

Psychological Bulletin

CONTENTS

GENERAL REVIEWS:

- Leadership, Membership and Organization:* RALPH M. STODDILL, 1.
The Indirect Assessment of Social Attitudes: DONALD T. CAMPBELL, 15.
An Experimental Analogue of Repression. I. Historical Summary:
 ANCHARD FREDERIC ZELLER, 39.

NOTE:

- On the Circularity of the Law of Effect:* PAUL E. MEEHL, 52.

BOOK REVIEWS:

- THORNDIKE's *Personnel selection*: S. R. WALLACE, 76.
 CHASE'S *The proper study of mankind*: H. J. MEYER, 78.
 SMITH'S *Psychological studies in twin differences*: ANNE ANASTASI, 80.
 JOHNSON'S *Essentials of psychology*: E. E. ANDERSON, 81.
 V. SCHILLER'S *Aufgabe der Psychologie*: F. A. PATTIE, 83.
 MACHOVER'S *Personality projection in the drawing of the human figure*: LIL-
 LIAN WALD KAY, 84.
 MAYO'S *Some notes on the psychology of Pierre Janet*: J. G. MILLER, 85.
 SYMONDS' *Dynaminic psychology*: F. J. SHAW, 86.
 THORPE AND KATZ'S *The psychology of abnormal behavior*: W. S. GREGORY, 87.
 FRANK'S *Projective methods*, and BELL'S *Projective techniques*: ARTHUR
 WEIDER, 89.
 WOLFF'S *Diagrams of the unconscious*: F. Y. BILLINGSLEA, 91.
 HARRIS' *Modern trends in psychological medicine*: LIVINGSTON WELCH, 93.

BOOKS AND MATERIALS RECEIVED: 95.

PUBLISHED BI-MONTHLY BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

EDITED BY

LYLE H. LANIER
New York University

CONSULTING EDITORS

STEGART H. BRITT
McCann-Erickson, Inc., New York
DORWIN CARY WRIGHT
University of Michigan
FRANK A. GELDARD
University of Virginia
JAMES J. GIBSON
Cornell University
DAVID A. GRANT
University of Wisconsin
WILLIAM T. HERON
University of Minnesota
ERNEST R. HILGARD
Stanford University
WILLIAM A. HUNT
Northwestern University

JUAN WALKER MACFARLANE
University of California
DONALD G. MARQUIS
University of Michigan
JOHN T. METCALF
University of Vermont
JAMES G. MILLER
University of Chicago
NEAL E. MILLER
Yale University
HILEN PEAK
Connecticut College
ROBERT W. SEARS
Harvard University
ROBERT L. THORNDIKE
Teachers College, Columbia University

The *Psychological Bulletin* contains evaluative reviews of the literature in various fields of psychology, methodological articles, critical notes, and book reviews. This JOURNAL does not publish reports of original research or original theoretical articles.

Editorial communications, manuscripts and book reviews should be sent to Lyle H. Lanier, New York University, New York 53, N. Y.

Preparation of articles for publication. Authors are strongly advised to follow the general directions given in the article by Anderson and Valentine, "The preparation of articles for publications in the journals of the American Psychological Association" (*Psychological Bulletin*, 1944, 41, 345-376). Special attention should be given to the section on the preparation of the bibliography (pp. 363-372), since this is a particular source of difficulty in long reviews of research literature. All copy must be double-spaced, including the bibliography.

Reprints. Fifty reprints are given, if requested, to contributors of articles, notes and special reviews. Five copies of the JOURNAL are supplied gratis to the authors of book reviews.

Business communications—including subscriptions, orders of back issues and changes of address—should be sent to the American Psychological Association, 1515 Massachusetts Avenue, N. W., Washington 5, D. C.

Annual subscription: \$7.00 (Foreign \$7.50). Single copies, \$1.25.

PUBLISHED BIL-MONTHLY BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

1515 Massachusetts Ave., N.W., Washington 5, D.C.

Entered as second class mail matter at the post office at Washington, D.C., under the act of March 3, 1879. Additional entry at the post office at Menasha, Wisconsin. Acceptance for mailing at special rate of postage provided for in Section 538, act of February 26, 1925, authorized August 1, 1947.

Copyright, 1950, by The American Psychological Association, Inc.

Psychological Bulletin

LEADERSHIP, MEMBERSHIP AND ORGANIZATION¹

RALPH M. STOGDILL

The Ohio State University

The present paper is concerned with a point of view regarding the relation of leadership to group organization. It represents one attempt within the Ohio State Leadership Studies² staff to clarify and systematize certain aspects of the leadership problem. Such clarification appears to be necessary as a preliminary step toward the development of an integrated program of research on leadership problems in formal organizations.

The pioneering work of Lewin (9), Moreno (10), and their followers has resulted in marked progress in the development of methods for studying leadership as a phenomenon of *groups*. However, comparable progress remains to be made in the development of methods for the study of leadership as an aspect of *organization*. Several factors appear to have operated as *barriers* to the development of scientific theory and method in this area. One is the lack of an adequate definition of leadership. A second is the fact that in much of the literature on leadership, the terms "group" and "organization" are used interchangeably or are defined in exactly the same terms. A third derives from two opposed theoretical approaches represented, on the one hand, by those theories of organization in which the leader is conceived as a symbol of authority or as an embodiment of superior personal traits, and on the other

¹ A cooperative contribution of the U. S. Navy, Office of Naval Research, and the Ohio State University Research Foundation. The opinions presented are those of the author, and should not be regarded as having the endorsement of the Department of the Navy. The paper was presented in part before the Midwestern Psychological Association Chicago, Ill., April 29, 1949.

² The Ohio State Leadership Studies are designed as a ten-year program of research on leadership problems in military, business, industrial, educational and civilian governmental organizations. The staff includes C. L. Shartle, Director; Alvin E. Coons, Melvin Seeman and Ralph M. Stogdill, Associate Directors; John Hemphill, Research Associate; Donald T. Campbell, Research Consultant; Richard T. Morris and Charles M. Westie, Research Assistants.

hand, by a type of group-oriented theory in which leadership appears to be regarded as a manifestation of social pathology. A fourth, and related obstacle, results from a reaction of social scientists against the authoritarian principles advanced in many discussions of organization. Some social theorists appear to reject all concepts of organization as authoritarian; and some researchers appear reluctant to deal experimentally with such concepts as responsibility, authority, stratification and similar phenomena related to organization. It is beyond the scope or purpose of this paper to portray the magnitude of the latter two difficulties. Nevertheless, it seems relevant to recognize the fact that they are present and act to the detriment of scientific work in the field.

The Ohio State Leadership Studies are being conducted on the basis of these assumptions: (1) that group organization is a recognizable social phenomenon in our culture; (2) that as such it is a legitimate subject for scientific study; and (3) that the variables of organization can be isolated and defined so as to permit their scientific study. It is the purpose of the present paper to examine various concepts relevant to leadership and organization, and to develop a formulation of the problem which will suggest hypotheses that can be subjected to experimental test.

GROUPS AND ORGANIZATIONS

Wilson (17) has reviewed the important sociological literature relating to concepts of the social group. He reports that in "current sociological literature one finds no consensus as to the meaning of the *group*," and concludes that much experimental work is yet to be done in order to delimit the group concept in any satisfactory manner. An important step in this direction has been made by Hemphill (5), who has devised scales for the measurement of such group dimensions as size, permeability, stability, viscosity, homogeneity of membership, and the like.

The most satisfactory definition available at the present time appears to be that of Smith (15) who defines a *social group* as "a unit consisting of a plural number of organisms (agents) who have collective perception of their unity and who have the ability to act/or are acting in a unitary manner toward the environment." Krech and Crutchfield (8) present a similar view. They state that "the criteria for establishing whether or not a given set of individuals constitutes a psychological group are mainly two: (1) all the members must exist as a group in the psychological field of each individual, i.e. be perceived and reacted to as a group; (2) the various members must be in dynamic interaction with one another."

A special kind of group is the *organization*. An organization may be defined as a social group in which the members are differentiated as to their responsibilities for the task of achieving a common goal.

Znaniecki (18) has reviewed the sociological literature relating to various concepts of organization. He stresses the fact that the terms *group* and *organization* are rather tenuous concepts, in that it is often difficult to determine whether a particular aggregate of persons constitutes a group, and that it may also be difficult at times to determine whether a particular group can be regarded as an organization. He points out that social organization

... can be realized only in a lasting "social group" or "association." Individuals belonging to such a group are aware that they will be regularly expected to perform certain actions, and some of them act as organizers, leaders, coordinators of the regular activities of others with reference to the common purpose. Not all of these individuals need be continuously active; indeed, in many groups a considerable proportion remain passive, acting only in reaction to the actions of others. The common purpose of the organized actions may be simple or complex.

Some of the consequences of distinguishing between the terms "group" and "organization" are the following: First, there is nothing in the term "group" which gives any clue as to the nature of leadership. Second, there is nothing in the group definition which provides any foundation for integrating leadership with group phenomena, except at a superficial level of social perception or interaction. Third, the group orientation can suggest research methods relating to leadership only in so far as the social group is defined in terms of organization. The concept of organization, however, with its implications for the differentiation of responsibility roles, does permit the study of leadership as an aspect of the relationships between members who are coordinating their efforts for the achievement of common goals.

A group may or may not have leaders. If it does have leaders, it is an organization, for at least some of the members are thereby differentiated from the others as to responsibility, or role expectation in relation to some common purpose. The members of a group may or may not have mutual responsibilities for a common task. If the members do have differentiated responsibilities in relation to common goals then the group is an organization—a particular kind of group. The continued presence of leaders and of responsibility differentiations in relation to group goals are indicative of organization. It may not always be easy to determine the exact point at which a group emerges into an organization.

LEADERSHIP AS AN ASPECT OF ORGANIZATION

The following definition of leadership may serve as a starting point for discussion. Leadership may be considered as the process (act) of influencing the activities of an organized group in its efforts toward goal setting and goal achievement. The definition of leadership relates it directly to the organized group and its goal. It would appear that the minimal social conditions which permit the existence of leadership are the following:

1. A group (of two or more persons).
2. A common task (or goal oriented activities).
3. Differentiation of responsibility (some of the members have different duties).

There are innumerable other group and situational factors which may influence leadership in varying degrees, but these appear to be the minimal conditions which will permit the emergence of leadership. There must be a group with a common task or objective, and at least one member must have responsibilities which differ from those of the other members. If all members perform exactly the same duties in exactly the same way there is no leadership. A leader then is a person who becomes differentiated from other members in terms of the influence he exerts upon the goal setting and goal achievement activities of the organization.

The foregoing discussion suggests that leadership cannot emerge unless the members of a group assume different responsibilities. It has been suggested that group organization also is founded upon differentiation of responsibility. It would then appear that leadership and organization are derived from a common factor or, viewed from a different light, that leadership is an aspect of group organization. This view has been expressed in various forms by writers in the field of business organization. Davis (3), for example, states that the

... development of organization structure is largely a problem in the division of responsibility, and involves two important problems: (1) the grouping of similar functions to form the various organization elements in a manner that will promote effective cooperation, and (2) the determination of the proper relationships between functional groups and organization elements, with a view to promoting both cooperation and effective executive leadership.

The definition of leadership does not specify how many leaders an organization shall have, nor whether the leadership influence of an individual is continuous or intermittent, nor whether the influence of the leader shall be for the welfare or detriment of the organization and its members. It merely specifies that leaders may be differentiated from

other members in terms of the extent to which they influence the activities of the organization in its efforts toward the achievement of goals. The definition of effective and ineffective leadership is an additional problem.

ASPECTS OF RESPONSIBILITY

Brown (2) in a challenging analysis of organization maintains that "An enterprise is a mosaic of its individual responsibilities. The sum of them must exactly equal the whole requirement of administration." He continues, "Responsibility is that part in administration which is assigned to a particular member of an enterprise. Its definition is an act of organization."

Responsibility cannot be regarded as a simple or uncomplicated variable. Jucius (7) writes,

By responsibility is meant, first, the obligation to do an assigned task, and, second, the obligation to someone for the assignment. But what is meant by obligation and how far does it extend? This implies a willingness to accept, for whatever rewards one may see in the situation, the burden of a given task and the risks which attend in the event of failure. Because of the rewards and penalties involved, it is highly essential to specify the limits of responsibility.

Formal organization can seldom define all the possible variations of responsibility and personal interaction to be expected of all members in all situations. Nevertheless, organization appears to be founded upon a basic system of stable expectations regarding differential responsibilities and relationships among the members. This is not a one-way process. That is, it is not the organization alone which sets up role expectations for its members. The members set up expectations for each other and for the organization as a whole. It is assumed for purposes of the present discussion that this principle applies not only to stratified organizations, such as military and industrial establishments. It applies as well to membership in any organized group, whether it be a business, political, educational, religious, fraternal, or social organization and regardless of size, stratification, purpose, or member characteristics. The essential relationship which makes possible the conduct of organized group activities is a differentiation of responsibility roles among the members. Without this there is no possibility of coordination or of leadership toward goal achievement. The very process of organization defines the responsibilities of the members and thereby the formal leadership of the group. It is true that in some organized groups, such as recreational groups, the responsibilities of members may appear to be vaguely defined. However, this is not equivalent to saying that no responsibilities exist.

Responsibility, in its broadest scope, defines not only the duties for which a member is accountable; it defines also the persons to whom and for whom he is accountable in the discharge of his duties. In doing so, it also defines a member's formal status, or location in the organization hierarchy. Authority and formal status systems in organization are but aspects of the division of responsibility.

Responsibilities in a systematic organization are determined by the assignment of persons to particular positions, the duties of which are outlined in an organization manual or organization chart. In less systematic organizations the responsibilities of a particular job or position may be determined by on-the-spot instructions, by general hints or by un verbalized assumptions. In a systematic organization an individual's *work patterns* (the tasks he *actually* performs) will correspond fairly closely with his *responsibility patterns* (the tasks he *is supposed* to perform). However, as the mission and activities of the organization change there will be found in many instances an increasing discrepancy between the tasks being performed and the responsibilities originally outlined and defined.

ATTRIBUTES OF ORGANIZATION

The studies of Roethlisberger (12) and others have directed attention to the factor of informal groups within formal organization. Informal organization, as usually defined, refers to the friendship groups and cliques—based upon close association, mutual interests or antagonisms, and the like—which develop within formal organization. It has been pointed out by Homans (6) that this conception is too narrow, since what is informal in a factory may be formal in a primitive society. Firey (4), who defines informal organization in terms of schism, presents a more useful approach to the problem. He maintains that “if we regard behavioral conformity, in which interactional processes are highly repetitive and synchronized, as the overt counterpart of a social system, then behavioral nonconformity may be taken as the overt counterpart of schism within a system.”

An organization in operation seldom corresponds exactly with the organization model as charted. The intervention of human social factors and other influences result in the emergence of informal organization, that is, in the development of work patterns and interaction patterns which do not correspond with responsibility patterns.

It would appear then that there are two fundamental sets of variables which define the operations of an organized group. These are:

1. *Variables which define formal organization.* These are:

- a. Responsibility variables (the work one is expected to do).

- b. Formal interaction variables (the persons with whom one is expected to work).
- 2. *Variables which define informal organization.* These are:
 - a. Work performance variables (the tasks one actually performs).
 - b. Informal interaction variables (the persons with whom one actually works).

If we regard the variables listed above as basic variables of organization, we can also regard them as variables of membership and of leadership. In other words, an organization can be studied in terms of these four types of variables: responsibilities, work performances, formal interactions and informal interactions. Leadership can also be studied in terms of the same variables.

Responsibility variables define the duties that the members are expected to perform. The responsibilities of a given position may remain the same, whether A or B occupies the position. *Work performance* variables are defined by the tasks performed and by the methods of their performance. Individual A may accept a position previously occupied by B. The responsibilities as defined by organization charts and manuals may remain the same, but the tasks actually performed by A may differ somewhat from those performed by B, and the methods of performance may vary markedly.

Formal interaction variables define the persons to whom and for whom the members are accountable, as well as others with whom they are expected to cooperate, in the discharge of their responsibilities. *Informal interaction* variables are defined by the persons with whom the members actually work and cooperate in the performance of their tasks.

Informal organization comes about as a result of the development of discrepancies (a) between work performance and responsibilities as defined and (b) between informal interactions and formally defined interactions. Thus leadership is ever confronted with the task of reconciling discrepancies—discrepancies between what ought to be done and what is being done, between goals and achievements, between organizational needs and available resources, between the needs of individual members and the requirements of organization, between formal lines of cooperation and informal patterns of cooperation.

An organization in action comprises a complex of many variables in interaction. In making a pictorial representation of a business organization, the usual procedure is to plot the division of formal responsibility on a two dimensional chart. The horizontal dimension of the chart shows the division of responsibility for various kinds of work. The vertical dimension of the chart shows the division of responsibility for different levels of decision-making, and indicates the persons to whom one is accountable and those for whose performance

one is accountable in the discharge of duties. This dimension defines the formal authority and status systems of the organization. Level (position in the organization hierarchy) and function (kind of work performed) are not independent dimensions. Although functions tend to differ from level to level, there is considerable overlap. Results from the Ohio State Leadership Studies (13, 16) have shown that the functions of top leadership tend to be supported at each lower level in the leadership structure by increasingly more detailed and routine work in the same functions.

Personal interaction can also be conceived as varying in both horizontal and vertical directions. The horizontal dimension is defined by the range (number) of members with whom an individual interacts. Some persons tend to work alone or with single individuals, while others are observed to work with large numbers of persons. The vertical aspect of personal interaction is defined by the number of strata (echelons) above and below his own in which a member works with others. Some persons may be observed to work only with others at the same level in the organization. Others tend to work only with subordinates, and still others tend to work only with superiors. These tendencies may or may not represent expression of individual differences in social interaction patterns. Results obtained thus far in the Ohio State Leadership Studies suggest that these patterns of interaction may be determined in part by the functions served by various types of positions. Technical consultants and staff aides tend to spend more time with superiors. Members in supervisory positions are observed, as would be expected, to spend more of their time with assistants, and subordinates. Members in coordinative positions tend to spend time with superiors and subordinates, as well as with associates at the same level in the organization. A member's function or duties may determine to a considerable degree which persons in the organization he may influence, as well as the nature of the influence that he can exert.

GROUP ORGANIZATION DEFINES AND DELIMITS LEADERSHIP

The very process of defining responsibility serves to structure and delimit the role that the leader may play in the organization. He cannot perform all the duties of all the members. His own accomplishment is therefore dependent upon the performance of others. His responsibilities are circumscribed by the outlined procedures and delegated responsibilities necessary for the achievement of stated goals.

Each member must work within the organizational framework which

defines the limits of his participation (how far he ought to go and beyond which he ought not to go) in performance of duties. It also sets the requirements for his cooperation with others and defines his relationships with his superiors and subordinates. This organizational structuring is not viewed alike by all persons. To some it appears as a barrier to participation or recognition. To others it appears as a prod and stimulus to greater effort and participation. For still others it provides a secure and comfortable sphere of activities and working relationships. Organization, therefore, in defining the responsibilities and working relationships of its members, sets up barriers to participation, as well as facilitating it.

Even as the organization sets boundaries by providing a framework within which members discharge their responsibilities, so the individual presents various barriers to the influence of the organization upon his own behavior and reactions. Some members may be limited in capacity to discharge their responsibilities, while others who are highly skilled in the techniques of their responsibilities are limited in capacity to interact with others. Each member carries into the organization his past experiences, his needs, ideals, personal goals, and commitments to other organizations, which may modify and determine his capacity for participation. It would appear that the extent to which the behavior of different members is determined by the characteristics of the group represents a continuum from little to great, and also, that the extent to which the behavior of the different individuals determines the behavior of groups may be conceived as representing a similar continuum.

It becomes apparent that a study in leadership represents a study of relationships, of variables in interaction. According to Pigors (11), a study of leadership must consider: (1) the leader, (2) the members as individuals, (3) the group as a functioning organization, and (4) the situation.

All organizations operate within a larger cultural and environmental framework. No organization can escape entirely the influence of the external situation. The organization may be influenced by the availability of resources, by changes in the social order of which it is a part, by competition of other organizations for the participation, resources or loyalty of its members, and by innumerable other factors outside the control of the organization itself. These factors also influence the leadership of the group.

LEADERSHIP AND EFFECTIVENESS OF ORGANIZATION

According to Barnard (1) the persistence of cooperation depends

upon two conditions: (a) effectiveness, the accomplishment of cooperative purpose, and (b) efficiency, the satisfaction of individual motives. Thus, although in many situations it may appear desirable to effect a maximum of goal achievement with a minimum of organizational expenditure, such a procedure might jeopardize the welfare or morale of the members. It then becomes evident that there are many situations in which organization is confronted by a complex of contradictory factors which must be considered in arriving at a decision. It also becomes apparent that the effectiveness of an organization cannot always be evaluated in terms of the degree to which it has attained its objectives. It may be necessary first to evaluate the goals and objectives themselves or the cost of their attainment. A carefully thought out discussion of factors to be considered in setting organizational goals, arriving at decisions, and evaluating the success of an organization has been presented by Simon (14). He states "The accomplishment by an administrative program of its organizational goals can be measured in terms of *adequacy* (the degree to which its goals have been reached) or of *efficiency* (the degree to which the goals have been reached relative to the available resources)." Simon, in agreement with Barnard, maintains that the criterion of adequacy alone is not valid as a measure of group accomplishment. He observes that "the fundamental criterion of administrative decision must be a criterion of *efficiency* rather than a criterion of *adequacy*. The task of administration is to maximize social values relative to limited resources."

If organizational goals are employed as reference points in evaluating effectiveness, then the goals themselves must be subject to evaluation. In addition, the cost (human or material) of goal attainment must be considered as a factor in evaluation. Both Barnard and Simon imply that organization cannot be regarded as a unit in isolation—or as a law unto itself. The motive of organization is the creation of social value or goods for its members, and these values bear some significant relation to the values of society in general.

Since leadership is related to the determination of group goals, it becomes apparent that the leader is seldom a free agent. In influencing the activities of the organization in its striving toward goal achievement he must consider certain social values, not only in relation to the members, but in relation to society as well. If he ignores the welfare of the members he is likely to lose their following. If he ignores the welfare of society he is likely to lead his group into difficulty. Thus leadership is subject to determination by factors which are external to the organization, as well as by internal group factors.

THE DEFINITION OF LEADERSHIP

The definition of leadership as a process of influencing the activities of an organized group in its task of goal setting and goal achievement should perhaps be reexamined. Does it define leadership? What are its implications? Admittedly, it defines only at a high level of generality. Certainly it does not include all social acts and influences, but it is nevertheless, an inclusive rather than a restrictive definition of leadership. Even so, it is more restrictive than most of those attempted in the recent literature. The definition restricts leadership to influence within the organized group. It does not imply domination or direction of others, nor any technique of influence; nor does it specify any particular member who should be regarded as a leader. The definition permits the study of any member of an organization to determine the extent of his leadership influence, and permits consideration of the possibility that every member may contribute toward determining the leadership of the organization.

The definition carries the implication that leadership may be not so much an attribute of individuals as an aspect of organization. In this respect it parallels the concept of authority. It is generally recognized that an executive in a business concern has authority in relation to his employees only during the time they are working as members of the organization. His authority does not extend outward into the direction of their personal or social lives. Nor does his position as an executive give him authority over other persons who are not members of his organization. In other words, authority is a relationship that exists between persons who are mutually participating as members of an organized enterprise. Authority is not an attribute of one or a few persons. Authority is an interactional process by means of which the organization defines for each individual the scope for action he has in making decisions, carrying out responsibilities, and enlisting the cooperation of others. The authority of any single individual will be largely circumscribed and defined by the authority of others, and at the same time, his own degree of authority will in part determine the authority of others.

Leadership appears also to be determined by a system of interrelationships. As such it must be regarded as an aspect of organization, just as authority is a derivative of organization. If leadership is determined by a system of interacting variables, then each of the several dimensions of responsibility and personal interaction might be conceived as representing a gradient of influence. If so, then it should be possible to measure leadership influences in terms of these dimensions.

Some members may be regarded as rating higher than others in leadership by virtue of the fact that they have responsibility for making decisions which exert a marked influence upon the activities of the organization. Some members may influence the activities of the organization as a result of personal interaction with other members, even though they do not hold positions of high level responsibility. Some members may rate high in both types of influence. It would not be expected that any organization could be found in which all influence is exerted by a single member. It would rather be expected that all the members of the organization could be ordered or ranked to some degree in terms of the influence they exert in various dimensions. The proposal to measure leadership in terms of the influence exerted by individuals may appear to contradict the statement that leadership is an aspect of organization rather than an attribute of individuals. But this is not a necessary conclusion. It was pointed out that authority is generally understood to be an aspect of organization. However, it can be observed that some members exercise more authority than others. The judgment can also be made that some persons have "too much" or "too little" authority. Such observations indicate an evaluation of conditions relative to various factors in the organization. In the same way it can be observed that member A exerts more leadership influence in some situations; while members B, C, and D exert more influence in determining activities of the organization in other instances. It may be that the leadership of A is circumscribed by the leadership of B, C, and D who are in competition with him; or it may be that the leadership of A is dependent upon the supporting leadership of B, C and D. In either event, the leadership influence of any one member is determined in part by the leadership exerted by others, and the balance may change from time to time.

SUMMARY

An organization is composed of individuals. Its existence is dependent upon the cooperation and performance of individuals who play different roles. Measures of authority, leadership, and the like, are but measures of aspects of organization, even though the measurements are made in terms of members and the relationships among members. Leadership exists only in so far as individuals, as members of organization, are differentiated as to the influence they exert upon the organization; and the leadership influence of any one member will be determined to a large degree by the total leadership structure of the organization. It is for this reason that leadership has been here defined in terms

of influence upon the activities of the organization, rather than in terms of influence upon persons.

The advantages of this formulation of the leadership problem are as follows: First, it removes leadership from the broad, vaguely defined realm of social interaction in general, and integrates it with the basic variables which describe an organized group. Second, and more important, is the fact that it suggests the development of methods for studying leadership as an aspect of work performance, work methods and working relationships.

An attempt is being made to develop such methods for the Ohio State Leadership Studies. For example, the goals and structure of organization and the responsibility patterns of members are determined by examining organization charts and manuals and by interviews with members of the organization. Work patterns are determined by modified job analysis procedures. Sociometric methods are employed to determine working relationships between the members and to chart the informal organization. The social values and role concepts of leaders and members are studied by means of attitude scales. These methods are supplemented by various check lists and rating scales.

In conclusion, a word of caution may be in order. The present paper has been concerned with a search for the minimal factors which will permit a functional integration of the concepts: leader, member and organization. In attempting to isolate these minimal common elements, many other important factors associated with leadership and group functioning have been excluded as not contributing to this central purpose. The present formulation represents merely one segment of a set of hypotheses to be subjected to experimental test.

BIBLIOGRAPHY

1. BARNARD, CHESTER I. *The functions of the executive*. Cambridge: Harvard Univ. Press, 1938.
2. BROWN, ALVIN. *Organization of industry*. New York: Prentice-Hall, 1947.
3. DAVIS, RALPH C. *Industrial organization and management*. New York: Harper, 1940.
4. FIREY, WALTER. Informal organization and the theory of schism. *Amer. sociol. Rev.*, 1948, 13, 15-24.
5. HEMPHILL, JOHN K. *Situational factors in leadership*. Ohio State Univ., Bur. Educ. Res. Monogr. 31, 1949.
6. HOMANS, G. C. A conceptual scheme for the study of social organization. *Amer. sociol. Rev.*, 1947, 12, 13-26.
7. JUCIUS, MICHAEL J. *Personnel management*. New York: Irwin, 1947.
8. KRECH, DAVID, & CRUTCHFIELD, RICHARD S. *Theory and problems of social psychology*. New York: McGraw-Hill, 1948.
9. LEWIN, KURT, LIPPITT, RONALD, & ESCALONA, SYBILLE K. *Studies in topological and vector psychology I*. Univ. Iowa Stud. Child Welf., 16(3), 1940.
10. MORENO, J. L. *Who shall survive?*

- Washington: Nervous and Mental Diseases Publ. Co., 1934.
11. FIGORS, PAUL. *Leadership or domination*. New York: Houghton Mifflin, 1935.
 12. ROETHLISBERGER, F. J., & DICKSON, WILLIAM J. *Management and the worker*. Cambridge: Harvard Univ. Press, 1939.
 13. SHARTLE, CARROLL L. Leadership and executive performance. *Personnel*, 1949, 25, 370-380.
 14. SIMON, HERBERT A. *Administrative behavior*. New York: Macmillan, 1947.
 15. SMITH, MAPHEUS. Social situation, social behavior, social group. *Psychol. Rev.*, 1945, 52, 224-229.
 16. STOGDILL, RALPH M., & SHARTLE, CARROLL L. Methods for determining patterns of leadership behavior in relation to organization structure and objectives. *J. appl. Psychol.*, 1948, 32, 286-291.
 17. WILSON, LOGAN. Sociography of groups. In G. Gurvitch and W. E. Moore (Eds.), *Twentieth Century Sociology*. New York: Philosophical Library, 1945. Pp. 139-171.
 18. ZNANIECKI, FLORIAN. Social organization and institutions. In G. Gurvitch and W. E. Moore (Eds.), *Twentieth Century Sociology*. New York: Philosophical Library, 1945. Pp. 172-217.

Received July 31, 1949.

THE INDIRECT ASSESSMENT OF SOCIAL ATTITUDES

DONALD T. CAMPBELL

The Ohio State University

In the problem of assessing social attitudes, there is a very real need for instruments which do not destroy the natural form of the attitude in the process of describing it. There are also situations in which one would like to assess "prejudice" without making respondents self-conscious or aware of the intent of the study. At the present time there are few if any indirect tests which could confidently be used for either of these purposes. There are, none the less, a considerable number of techniques that have been partially explored and validated. It is the purpose of this paper to survey such techniques and to present a point of view with regard to the problem of indirect measurement.

Current interest on the part of social psychologists in the indirect assessment of attitudes is perhaps primarily an aspect of the larger projective test movement in personality study. However, as will be seen in the course of this survey, there has been an interest in this approach from the very first efforts in attitude measurement (81), anticipating by some ten years the current interest in "projective techniques."

The terms "indirect" and "projective" have been used to refer to both *disguised*, and to *non-structured* measurement efforts. Using these two terms alone, one could distinguish four types of tests:

1. *Non-disguised-structured*: the classic direct attitude tests of Thurstone (78), Likert (49), *et al.*
2. *Non-disguised-non-structured*: the free-response interview and questionnaire approaches, the biographical and essay studies.
3. *Disguised-non-structured*: the typical "projective" techniques.
4. *Disguised-structured*: tests which approximate the objective testing of attitudes.

It is with the latter two categories that the present review is concerned, although some items that properly belong in category 2 will be included where they represent deliberate efforts at "projective" attitude testing.

While a formal division of content on these two criteria suffices, the writer will, in discussion, use a third which overlaps but does not duplicate the other two. This third criterion is that of dependence upon voluntary self-description as opposed to diagnosis based upon differential performance in an objective task. Upon this latter criterion rather than "structuredness" will rest the primary distinction between our two types of methods.

DISGUISED-NON-STRUCTURED TESTS OF SOCIAL ATTITUDES

In this category are included those techniques which offer the respondent opportunity for the spontaneous expression of attitudes in an ambiguous or non-structured setting. Most of these techniques are borrowed quite directly from well established clinical tools, and will be classified on that basis.

Approaches based on the Thematic Apperception Test. Though by no means the first, the indirect test most widely cited in previous surveys of the literature (18, 42, 54, 82) is that of Proshansky (65). He intermingled ambiguous pictures of labor situations with the more usual T.A.T. scenes. The pictures were presented to a group by means of slides, with each person being asked to write for two and one-half minutes on what he thought the slide represented. Slides were shown for only five seconds. Proshansky found that ratings made from the resulting descriptions correlated .77 and .67 with a direct verbal scale of attitudes toward labor. A more elaborate proposal for the use of a similar technique with attitudes toward Negroes and Jews has been made (51), but as yet, no results are available.

Specially designed Thematic Apperception pictures have also been used by Frenkel-Brunswik, Sanford and Levinson in their extensive study of the personality correlates of prejudice (24). In this research the purpose was not so much to measure prejudice as to get a more detailed and qualitative picture of its expression. Nevertheless, the complicated interrelationships they found qualify the use of such pictures as attitude measuring instruments. For example, many prejudiced women told warmer and more sympathetic stories to a picture of a Negro mammy than they did to pictures of an elderly white woman (23). While such a finding is consistent with personality theory, it points to the danger of an over-simplified one-to-one interpretation of such material.

Essentially identical in technique is the "Human Relations Test" developed by the staff of the College Study in Intergroup Relations (15). This consists of ten drawings, each portraying an ambiguous intergroup contact. For example, a crowded street-car scene and a basketball scene are used. The respondents are asked to write a short story about each scene, telling what happened, how it came about, and what will happen next. The test is available on printed cards or projection slides, and since it is being used by a number of cooperating colleges, and at a wide variety of age levels, there is some promise of a standardized test emerging.

Johnson (39) has successfully studied the development of Anglo-Spanish attitudes in the Southwest through stories told by children to a specially designed series of pictures. He used six carefully selected conflict situations. These were duplicated in three forms: with all Anglo characters (for use with English-American children), with all Spanish characters (for use with Spanish children), and with mixed Anglo-Spanish characters (for use with both groups). Attitudes were assessed by contrasting responses to identical situations when they involved Anglo-Spanish conflict and when depicted by their own group members only. Quantification was achieved by having two judges categorize individual responses on a number of dimensions. Reliability coefficients for six sub-groups were over .90.

Approaches utilizing doll play techniques. In the first explicit attempt to use a projective technique in the assessment of social attitudes, Dubin (19) utilized toys in a fashion similar to the "play techniques" of the clinical psychologist. Using ten adult respondents, he asked them to "construct on this table a dramatic scene or scenes of the world as you see it today" and later "Now make a dramatic scene or scenes of the world as you would like it to be." From report and pictures of these scenes, three judges were able to estimate answers on 21 direct attitude questions dealing with labor, the Negro, internationalism, etc. with an average rank order correlation of .49.

For use with children, Arnold Meier has developed an interviewing aid, the "What Would You Do?" test, which is "projective" in the present sense. Doll cut-outs including minority children are manipulated on background drawings which depict such situations as entering the home, school scenes, etc. (15).

Similar to these approaches is the "movie story game" developed by Evans and Chein (21) for use with children from eight to twelve. On a miniature stage, Negro and white dolls are manipulated, with the child being asked at various points what the identified doll would say. Particular attention is paid to the patterning of segregation responses with acquaintances and with strangers. Preliminary studies indicate effective disguise and general meaningfulness for the test.

Hartley and Schwartz (33) have combined doll-play with pictorial material in the investigation of intergroup attitudes of children five to seven years of age. Montage background compositions carried characteristic symbols of Jewish religion in one case, Catholic in another, and in a third indicated typical middle class surroundings with no indication of religion. Identical family sets of dolls were placed upon these three backgrounds, and all were available for use by the child in playing out situations, such as a birthday party, school bus, meals, etc. In the preliminary report it is indicated that children identify with some accuracy the religious symbols, and that their play indicates intergroup attitudes in a meaningful way.

Sentence completion tests. At the Ohio State University, Shirley Wilcox Brown is using a modification of Rotter's test (68) in investigating attitudes toward the Negro. A 40-item schedule has been prepared, in which are imbedded some 20 sentence fragments dealing specifically with the Negro problem. Examples of relevant and neutral items are as follows:

1. I feel . . .
2. Skin color . . .
3. I hate . . .
4. Maybe . . .
5. Some lynchings . . .
6. The K.K.K. . . .
7. It seems to me that segregation . . .
15. Negro body odor . . .
37. Racial intermarriage . . .

In this test, no real effort at dissembling is made. The cooperating respondent is asked to volunteer a description of his own attitudes, with no more disguise than in the direct tests. The instructions read "Complete these sentences to express *your real feelings*. Try to do every one. Be sure to make a complete sentence!" Preliminary applications of the test indicate its power to discrimi-

nate between criteria groups. Scoring has been done by both coding and rating procedures. The indications are that the test can provide a reliable measure of attitudes.

Another twist to the sentence completion notion is being used in the study of personality and prejudice among school children conducted by Else Frenkel-Brunswick and Harold E. Jones (22). Here the test avoids all mention of minority groups, but provides stereotypic statements which may be completed with names of various minority groups or others. Examples of items are:

Are there some people who are mean? WHAT PEOPLE?

It would be better if more of a certain type of people were allowed to come into the United States. WHAT PEOPLE?

Some people are poor and it is their own fault. WHAT PEOPLE?

This test has elicited mention of foreign and minority groups from about one third of the children to which it has been administered. From another third or so come one or more anti-prejudice statements. A portion of the students make no responses classifiable in either way, and are thus not effectively evaluated. Using a net score (subtracting anti-prejudice responses from the total of prejudiced ones) corrected reliability figures run around .6 to .8 and correlations with a highly reliable direct test are on the order of .5. The approach is most satisfactory for the comparison of groups of respondents, and for the evaluation of the relative salience and extremity of attitudes toward different minority groups. In addition, unique data on the uniformity of stereotyping are provided.

While these two sentence completion tests make a deliberate effort to achieve a projective or indirect character, they are not too different from any free-response questionnaire on attitudes. Compare, for example, Zeligs' approach in which school children were asked to "write the most interesting true sentence" they knew about each group, within a one minute limit (84).

Miscellaneous non-structured techniques. The Rosenzweig Picture Frustration Test has also been modified for the measurement of social attitudes, a task to which it is obviously appropriate. This test requires the respondent to fill in the balloons in a series of cartoon drawings involving frustrating face-to-face inter-group contacts (6).

Probably belonging within the limits of our general category is the study by Fromme (25). He presented to the respondents five political cartoons, each with four alternative captions, covering a wide range of pro and con opinion. The respondent was asked to pick the best caption, and this choice, plus the discussion resulting, was utilized in a qualitative analysis of attitude structure.

The procedures just discussed are all indirect and "projective" in current usage of the terms. At the risk of some oversimplification, the assets they have in common may be emphasized and the stage set for contrasting them with the second group of procedures.

While tests such as these may be in part disguised, in that the experimenter does not tell the respondent his real purpose and may indeed substitute a false justification, their primary asset is that of securing an expression of attitudes in a more natural and spontaneous form, allowing opportunity for each individual to "project" upon a

neutral screen his own integration of the problem. In contrast, the usual direct paper-and-pencil attitude test, requiring the endorsement of prepared items, may be said to force artificially the expression of attitudes into a preconceived and common mold. The great advantage of these non-structured tests is one of freedom, rather than of disguise. While some respondents may complete their assignments unaware of the experimenter's interest, in a tense situation one could hardly use any of the above devices expecting to get unconscious or uncensored expressions of attitude from uncooperative respondents. Essentially these tests are "voluntary" in the same sense that the usual attitude test, interest inventory, or neuroticism test is. The respondent is told (either directly or in effect) that there are no right or wrong answers, and he is placed in a situation in which a voluntary and arbitrary performance is acceptable.

DISGUISED-STRUCTURED TESTS OF SOCIAL ATTITUDES

The approaches which we will consider below differ from the ones mentioned above perhaps only in degree or in relative emphasis. Yet the distinction involved is important. The characteristics of these disguised, non-voluntary tests, can be stated in a number of ways:

The respondent participates in an objective task, in which he seeks right answers. The voluntarism of the usual projective techniques is lacking. To the respondent, the situation is similar to that of an achievement or ability test.

All respondents have a common motivation in taking the test. All, we may assume, are seeking to perform adequately on the same objective task. Attention is focused on a common goal, oblique to the experimenter's purpose.

Rather than stressing the freedom and lack of structuring, there is an attempt to diagnose attitudes from systematic bias in the performance of an objective task. The test may be highly structured—directly scorable—and still offer opportunity for unconsciously operating bias to distort behavior in a systematic and diagnosable manner.

Here is a simple formula for constructing such a test. Find a task which all your respondents will take as objective, and in which all will strive to do well. Make the task sufficiently difficult, or use a content area in which the respondents have had little experience or opportunity for reality testing. Load the test with content relative to the attitude you study. Look among the responses for systematic error, or for any persistent selectivity of performance. If such be found, it seems an adequate basis for the inference of an attitude.

To be sure, this approach has to some extent been used in the projective techniques discussed above. The T.A.T., for example, is introduced as a "test of creative imagination" (58), but little attempt is made to carry out effectively this pretense pattern. Respondents accept

the task with a wide diversity of motivations, varying levels of seriousness, curiosity, and suspiciousness. And while all may be "projecting," the motivational situation is highly un-uniform, and meaningful comparisons between the responses of the serious, the suspicious, and the playful can hardly be made. Similarly, the "play technique" has been given to college men with the excuse that the experimenter was interested in "ideas for moving picture plays" and wanted the respondent to construct a "dramatic scene" (58, p. 553); and the sentence completion test has been given as an ability test (as Ebbinghaus intended it), with instructions to complete the sentences in any grammatical fashion as quickly as possible. But the disguise in these typical projective tests seems to be half-hearted, or implausible, weakened by the explicit freedom of response allowed. At best, the respondents accept the task as meaningless, or as a psychologist's mystery. It is probably common that something of the intent of the experimenter is divined, and the task completed despite this.

Information tests. To make the model of the indirect, disguised, non-voluntary test of social attitudes more explicit, let us examine the use of an information test for attitude diagnosis. The reader has probably commented at one time or another upon the inextricable interrelationship between people's attitudes and what they take to be facts. You would all probably agree that in a detailed test of information, the direction of people's guesses or misconceptions will frequently bear a relationship to their attitudes. In a complementary fashion, a given person's knowledge is apt to reflect in its unevenness his selective awareness and retention, or his biased sources of information. Newcomb has dramatically portrayed the non-random character of right and wrong answers on an information test in his study of "The Effect of Social Climate upon some Determinants of Information" (62). While selective awareness and divergent sources of information were primarily involved, Newcomb comments with regard to the difficult information items "the direction of guessing is altogether likely to be weighted toward the subject's attitude. If this reasoning is correct, the . . . test tends to become itself an attitude test." Coffin (14) and Smith (76) have found beliefs of factual type statements to correlate highly with related attitude tests. Indeed, in looking over the high correlations found in the literature on the relationship between information and attitudes, one is tempted to guess that some re-interpretation of them is in order. Is it not likely that many items were such that persons of liberal, tolerant attitudes found guessing easier?

Loebowitz-Lennard and Riessman have proposed that an information test be used to measure indirectly attitudes toward Negroes and Jews (50). Hammond (31) has actually used such a test to measure attitudes toward "labor-management" and "Russia," and his work is worthy of some detailed comment here. We have mentioned that not only guessing behavior but also differential patterns of information may be diagnostic of attitudes. Hammond eliminated the latter by the "error-choice" technique, in which the respondent was forced to choose between two alternative answers each of which was, by intent, equally wrong, but in opposite directions from the correct answer. Such items were:

"Average weekly wage of the war worker in 1945 was (1) \$37.00, (2) \$57.00." "Financial reports show that out of every dollar (1) 16¢, (2) 3¢, is profit." Scoring these information guesses as attitude items gave total scores on 20 such items that differentiated a labor union group from two business clubs with almost no overlap. In spite of the small number of cases (18 union and 42 business) the critical ratios were 11.3 on the Labor-Management, and 12.5 on the Russia questionnaire. Reliabilities on the two scales were roughly estimated at .78 and .87 respectively.

What we have here is an objective test situation (which could be made more so by using more alternatives and including a correct one) in which people's errors are not random, but systematic. The presence of biased performance clearly necessitates the inference of some underlying process, which we choose to call attitude. The claim for face validity on such a test seems to be stronger than the one that can be made for either the direct or unstructured attitude test. Is not systematically biased performance in dealing with environmental actualities the essential practical meaning of attitude?

Parrish (64) has more recently used a regular multiple choice information test as an indirect indicator of attitudes, in conjunction with a direct attitude scale and a second indirect test involving estimating opinion poll results (see below). All three tests effectively reflected 12 minutes worth of anti-Kuomintang propaganda in approximately the same degree when the greater reliability of the direct test is considered. The pseudo-information test had 32 items and a reliability of .66 (using Kuder-Richardson formula 14 modified for multiple levels of response). It correlated .67 (corrected for attenuation) with the direct test (10 items, reliability .91) given at the same time (both anonymously, so there is no reason to suspect the direct test). The correlation with the public opinion estimate test was .59.

Kremen (43) attempted to evaluate the effect of student role-playing of a discrimination episode upon attitudes toward the Negro, using a multiple-choice information-type indirect test and a direct test. While neither test reflected the role playing in mean scores, her findings have importance for attitude measurement, inasmuch as role-playing *lowered* the relationship between the direct and indirect test. The results from ten sociology classes, (five pairs in which the instructor and subject were the same) are presented below:

	<i>Correlation between Direct and Indirect Tests</i>				
<i>Role-Playing Classes:</i>	.39	.36	.20	.53	.38
<i>Paired Control Classes:</i>	.54	.70	.35	.58	.60

For all classes combined, the direct test had a reliability of .89, the indirect test .42 (Kuder-Richardson, Formula 14). The explanation of this phenomenon is not obvious from her data insofar as analysed. She does not report reliability values separately for the ten classes, although this step is planned in further analysis of the data.

R. B. Cattell (12, 13) has undertaken a large scale exploration into the "objective" measurement of attitudes, using a rationale very similar to the one presented here for non-voluntary attitude tests. In this study, a large number of ingenious methods have been tried, and others suggested.¹ The methods have

¹ Omitted from the present review are a number of physiological and related psychological measures, such as P.G.R., writing pressure, pulse pressure, speed of reading, fluency, and speed of decision, which Cattell also investigated.

been applied to a number of attitude-interests, such as playing more sociable games, excelling in one's career, and being smartly dressed, which differ greatly from the social-problem topics of most of the other studies reviewed. Two of the methods may be classified with information test approaches. His measure of "False Belief (Delusion)" (12) was presented to respondents as a multiple-choice information test. A sample item: "During the war church attendance increased greatly and, since V-J day it has: (declined slightly; tended to increase still more; stayed at its high peak; returned to its pre-war level; fallen to its lowest point since 1920)." Presumably the different responses were given different weights as in a direct test and as done by Parrish and by Kremen. Applied to a religious attitude, ten such items gave a Spearman-Brown reliability of .53, and correlated .33 with a paired-comparisons preference test, and .10 with records kept of time and money expenditure.

Cattell's other information test approach was more straight-forward. As mentioned above, we may expect that a person will reveal his attitudes not only in the direction of his "errors" but also in the selective character of the "right answers" he knows.

Cattell prepared 10-item information tests for each of a number of attitudes. The items dealt with the "knowledge required in following the course of action connected with the attitude" (12). For twelve topics, these tests, scored in terms of the number of items correct, provided Spearman-Brown reliabilities ranging from .13 to .92, with an average of .59. Correlation with records of time expended on the various activities averaged but .16. With money spent on activities the average value was .15, and with paired-comparison preferences, .16. In this approach, Cattell has resurrected an approach to the "objective" or "indirect" measurement of vocational interests which flourished during and after the first world war. Fryer (26) and others provide summaries of such research. No longer called "interest" tests, these selective vocabulary and object recognition tests are today firmly established among vocational selection and guidance tools.

Estimation of group opinion and social norms. As Travers (79), Wallen (80), and others have demonstrated, there is persistent correlation between a person's own attitude and his estimate of group opinion. While a few persons may chronically underestimate the popularity of their own opinions, the prevailing tendency is to overestimate the size of the group agreeing with one's self. Biserual correlations between endorsement of a dichotomous item and the estimation of the proportion of the group endorsing show few values below .2 and characteristically run from .4 to .6 or higher. This relationship seems to indicate the possibility of assessing a person's attitudes by giving him the entirely objective task of estimating group opinion. If one regards correlation with direct attitude expression as validation, one starts with very high single item validities indeed. If from a single item, one can predict a person's attitudes to the extent of a correlation of .40, how well might one not be able to predict from twenty such items? In several attempts along this line at the Ohio State University, conducted under the direction of the present author, the results have fallen considerably below such optimistic hopes. From combined estimates of group opinion, the best correlation obtained with a direct attitude score based on the same items runs around .60, not better than the best single item-estimate correlation. More detailed exploration of the technique is being made, how-

ever. Parrish (64) used a public opinion estimate test, with items dealing with U. S. opinion on China, the opinions of Americans with experience in China, and the opinions of the Chinese. Only the first type are directly comparable to the Travers and Wallen situation, and these did not come out the best in the item analyses. The test, containing only 10 items, had a Kuder-Richardson reliability of .53, correlating .67 with the direct test when corrected for attenuation. In this instance there was no direct duplication of content in the direct attitude test items and in the opinion estimates.

The technique is not entirely new. In 1929, Sweet, using a test designed by Goodwin Watson, found boys' estimates of group opinions valuable in diagnosing adjustment problems (77). More recently, Katz has reported results confirming the general relationship between own attitudes and estimates of group opinion, although he does not give correlation coefficients (41). As part of Section A in their "Sentiments" examination, Murray and Morgan (59) ask the respondent to guess what the majority of people believes or prefers, e.g. "what are the three most popular things to do?" Newcomb (61) made one of the first uses of the percentage estimate of group opinion, but has not reported his results in terms of the responses of individuals. It might be pointed out that the correlation involved in these studies is not necessarily to be interpreted as projection. In a group of any size, the respondent has an uneven acquaintanceship, and probably associates more with those of tastes like his own. Basing his estimates of group opinion upon his own experience in the group, his error may be in part a "sampling bias" as well as biased perception.

F. C. Bartlett (4, 5) and D. M. Carmichael (9) before the war were experimenting with the use of predictions of social events as diagnostic of attitudes. The respondent is given a situation and some background information, and then asked to make a prediction. Predictions regarding the likelihood of cooperation between social groups have been used, for example. The approach can probably be presented as a test of ability, and would seem to offer the possibility of wide variation in content. McGregor (53) has also pointed out the relationship between predictions and personal attitudes. Using predictions of another kind, Davis (17) has devised a simple test for use with children which could easily be given in an objective frame of reference, and which already has the desirable quality of asking the respondent to talk about somebody else rather than himself. (Indeed, his technique is practically identical with the first "projective" personality tests designed by Cattell (10) on the model of multiple-choice intelligence test items.) In Davis' test, situations are presented, and the respondent predicts the protagonist's action by selecting from several prepared alternatives.

Noland (63) used a projective approach toward estimating social norms with one set of items in his extensive study of absenteeism, asking "how important a cause of absenteeism do you think each of the following." These single items predicted absenteeism with a correlation of .50 or better. The value of the finding is weakened, however, by having as a criterion of absenteeism voluntary reports on the same (unsigned) questionnaire.

Tests of ability to do critical thinking. During the last three or four years of his life, John J. B. Morgan had been working on another technique that admirably meets our criteria (55, 56, 57). He used a syllogisms test, with a large number of possible conclusions provided, from which the respondent was to

pick the logically correct one. Following Sells' study of atmosphere effect, the syllogisms were invalid. In spite of the invalidity of the syllogism, college students will concentrate their choices on the one or two wrong alternatives favored by the atmosphere effect. To use this phenomenon in diagnosing attitudes, Morgan gave the same syllogisms in both impersonal or abstract form (e.g., "No A's are B's. Some C's are B's. From these statements it is logical to conclude: No C's are A's, etc.") and with content ("A trustworthy man does not engage in deceitful acts. The bombing of Pearl Harbor by the Japanese was a deceitful act. From these statements it is logical to conclude: No Japanese are trustworthy, . . . etc."). By studying the shift in the popular responses from the nonsense to the meaningful form, Morgan attempted to diagnose group attitudes. While the technique needs much further study, it seems well worth such elaboration. In particular, an attempt should be made to control the effect of greater tangibility, per se, and also the intrusion of evidence from outside the syllogism which might be confounded with purely attitudinal effects. Note that in this particular study, the shifts of the group showed more tolerance for the Japanese than otherwise. Lansdell has utilized the technique for studying covert attitudes toward marriage, attempting to apply it for the diagnosis of the attitudes of a single individual (45).

Ruch, in his 1937 edition of *Psychology and Life* (69, pp. 633-634), reports briefly upon a classroom experiment in which preferences for ethnic groups were reflected in the ability to judge the validity of syllogisms about the same groups. In this instance, no choice of conclusions was offered, but rather the students judged the validity of the complete syllogism as stated. Lefford (46) has shown in more detail the disrupting influence of emotional subject matter upon such judgments.

The syllogism approach, and a number of others, have been anticipated by Goodwin Watson's *Measurement of Fair Mindedness*, published in 1925 (81). By fair-mindedness Watson seems to have meant something roughly equivalent to tolerance, critical thinking, or open-mindedness. His six sub-tests were designed to provide scores on this trait, which they did with adequate reliability. But Watson recognized that in the "errors" or lapses from fair-mindedness were clues as to the biases or attitudes of the respondents, and a secondary "analytic" scoring was made in these terms. Three of the sub-tests we would characterize as being direct² and of the others, we will classify two under the heading of "critical thinking," treating the third—*Moral Judgments*—below. The *Inference Test* provided statements of fact, followed by several conclusions that might be drawn, only one of which is logically justified, the others offering opportunities for the intrusion of personal biases in various directions. This test has been expanded in the Watson-Glaser Tests of Critical Thinking (28) which could be scored for various prejudices although they have not been so

² These are the *Word Cross-out Test* in which the respondent crosses out "distasteful" words, as in Pressey's test, the *Degree of Truth Test* allowing for extreme or moderate statements of belief and disbelief and, the *Generalization Test* in which a number of statements about various groups are presented, and the respondent is to indicate whether the statement is true of "All, Most, Many, Few, or No" members of the group. (This is probably one of the first tests of "stereotyping." It might be classified as indirect if administered as an information test.)

used as far as the present writer is aware. Gilbert (27) has made use of this same technique in a study in which high-school children were given hypothetical racial problems for which they could choose strictly logical conclusions or conclusions showing bias.

In Watson's *Arguments Test* both pro and con arguments on various topics are presented, and the respondent evaluates these for their strength as arguments. The test taps the tendency for a person to view as strong those arguments which support his own point of view. Robbins (66) has recently confirmed the principles underlying this test, although his data are not in such form as to indicate whether or not accurate inferences as to attitudes could be made from such judgments. He found that students performed better in judging the adequacy of reasons for points of view with which they agreed than those with which they disagreed.

Tests employing bias in perception and memory. The demonstrations of systematic attitudinal bias in highly "objective" perceptual and learning tasks have provided some of the most important developments of the past 15 years for attitude theory (2, 20, 48, 73, 74). The use of errors in seeing and remembering in the diagnosis of individual attitudes epitomizes the disguised, structured, and non-voluntary assessment of attitudes.

Pioneers in the use of this approach are E. L. and R. E. Horowitz (37). In their study of the development of social attitudes among the children of a southern community they invented a number of techniques which have merited a great deal more re-use than they have seen. While designed to portray group differences, many of them should be appropriate to individual testing. Their *Aussage Test* involved exposing a complicated picture for two or three seconds, following which they tested for perception and memory through a series of standardized deliberately leading questions. For instance, the question "who is cleaning the grounds" to Picture 10 brings answers referring to a non-existent Negro in some 70% of the cases. To Picture 4, the misleading question "What is the colored man in the corner doing" brings a steadily increasing proportion (in reports from higher age groups) of menial activities for the non-existent Negro man.

Their *Perception-Span Test* involved a series of posters each having pictures of 10 items mounted upon it. These were exposed for 10 seconds, and the children asked to "tell all the pictures you can remember." While there were difficulties with perseveration and blocking, some age group differences were noted, with the younger children failing to note Negroes, and the older ones showing a selective awareness for them. A *Recall Test*, asking for the reproduction of ten words used earlier in a word association test, seemed also to indicate that in the younger years the word "Negro" was less well remembered than would be expected. The *Pictorial Recognition Test* involving the recognition of faces previously exposed, finds the faces of Whites more frequently recognized than the faces of Negroes.

Seeleman's study (71) confirms this finding with respect to individual differences within the group. Presenting sets of Negro and white photographs she tested for later recognition by asking that the previously exposed pictures be picked from a larger group. The discrepancy between memory for white and Negro faces was checked against attitudes on a direct test. Using people in the extreme quartiles on the direct test only, the test and memory-bias scores

correlated .64 and .71 with different populations. In a paired-associates memory test involving complimentary phrases to be paired with white and Negro pictures, error discrepancy scores showed no correlation with the direct test, but the tendency to assign favorable phrases to the Negro (regardless of correctness) correlated .66 with the direct test in the one population for which the correlation is reported. These values are high enough to suggest that with refinement an indirect test might be achieved.

On a preliminary experiment, Murray Jarvik (22) attempted to utilize memory distortion as an indicator of attitudes. Sixth, seventh and eighth grade students in a rural school were read five minute stories and then were asked to write down all they could remember of the story. The stories had simple dramatic plots, but were full of confusing detail and lacunae with regard to specific names, characteristics, and ethnic identities. Opportunity was given for memory distortion in the direction of common stereotypes. The results were essentially negative. However, motivation and literacy were low, and the task was a part of a rather long testing program. Note also that memory distortion usually occurs over longer spans of time, and that had the test employed selective retention rather than distortion it might have been more successful.

Cattell (12, 13) utilized three techniques involving memory and perceptual distortion. His test of "Immediate Memory" was based upon selective recall for attitude relevant statements from sets of twelve presented at one-second intervals. In the total, there were over 500 statements. Spearman-Brown reliability figures for eleven attitudes ranged from .13 to .86, averaging .50. Correlations with other measures were essentially zero (with time expenditure, average .01; with money expenditure, average .03; with paired-comparisons preferences, average .13; with the information test (see above), .01.). In the scoring, both "facilitating" and "frustrating" statements relative to the attitude were pooled. Cattell recommends that in future research these be separated.

Cattell's "Distraction" method involved the exposure for ten seconds of statements related to the attitudes. Around the statement were scattered twelve or thirteen nonsense syllables. Subjects were held responsible for recalling the statement and the nonsense syllables. Scoring was in terms of decrement in nonsense syllable recall for statements related to a given attitude. The original hypothesis was that the stronger the attitude, the poorer the nonsense syllable learning, through distraction. Actually, the opposite effect was found. With a Spearman-Brown reliability of .64 for one attitude (being smartly dressed), the correlation with the paired-comparison preference measure was $-.29$, with time expenditure, $-.10$, with money expenditure, $-.08$, and with the information test (discussed above) $-.35$. Since only ten items were involved, these values are regarded as sufficiently high to encourage further work with the method. The "Misperception" method involved the tachistoscopic exposure of attitude statements for one second, with the respondent responsible for repeating the statement and noting the misspellings. The overlooking of misspelling was regarded as a sign of strong attitude. Ten items dealing with the desire to play more sociable games provided a Spearman-Brown reliability of .43, and correlations with the paired-comparisons preference measure of .00, the expenditure of money of $-.01$, and the information test of $-.13$.

These methods of selective memory and selective distractibility have been anticipated in the indirect measurement of occupational interests by Burt in 1923 (7, 8, pp. 334-343). In a paired-associates memory test, one half of the pairs were relevant to agricultural engineering. The ratio of memory for these pairs as opposed to memory for other pairs correlated .30 with instructors ratings as to interest and industry. In a cancellation test, in which irrelevant words were to be crossed out, different paragraph contents were employed. Differentials in efficiency of cancellation correlated .30 with the above-mentioned instructors ratings. Contrary to Cattell's findings, the greater distractibility was associated with stronger interest.

Tests involving ability to judge character. Murphy and Likert, in *Public Opinion and the Individual* (60), utilized a wealth of techniques anticipating the projective testing movement. Among these was the "Ratings from Photographs." Following the general framework of Rice's classic study on stereotypes, they provided labeled pictures of a union president, a railroad magnate, a pacifist, a Negro champion of Negro rights, etc. Respondents were asked to judge from these labeled photographs the character of the pictured person, in terms of courage, selfishness, intelligence, conceit, sympathy, practicality, and sentimentality. Contrary to expectation, they found no relationship between attitudes as measured in a variety of paper and pencil tests and these picture ratings. However, since the making of character judgments from photographs comes as near to being a paradigm of psychiatric projection as can be found in present test approaches, it might be suggested that this technique be further explored before being abandoned. In line with the tenor of the present discussion would be the suggestion that the objectivity or "pseudo-objectivity" of the task be increased, with the use of some such label as "social intelligence test" or the like.

In precisely this model is the "Faces Game" originally developed by Marion Radke (67, p. 50-51) and being used as modified by Isidor Chein and Iljana Schreiber in research of the Commission on Community Interrelations of the American Jewish Congress (16). In this test children are asked "Let's see how good you are at telling what people are like just by looking at their faces." Thirty four sets of four pictures (two white and two Negro for all but four sets) are presented, each with a question such as "one of the girls in this row is very lazy and never bothers to do anything" or "one of the boys in this row is the best sport in his class." Two scores are provided: one the "white salience" score (the total mentions of white children for items either good or bad), and the *prejudice score* (the ratio of unfavorable choices of faces of the other group to total choices of other groups). The test produces significant individual and group differences, and showed retest reliabilities over a six month period of .50 and .36 for white and Negro children, respectively, on the prejudice score. For the white salience score, the values were .32 and .16. The researchers do not regard the test as satisfactory for general use as yet.

Hsü (38) had three women graduate students sort male photographs for handsomeness, and ten days later for judged membership in the Communist party. The correlations were negative (-.50 and -.21) for the two women who were anti-communist and positive (.51) for the one woman who was relatively pro-communist. After reading a *Time* report on the blockade of Berlin, the correlations for sorts on a second set of the photographs were negative for all three women. While this study does not provide in itself an adequate basis for

measuring attitudes, it does hold out promise for more extensive work with character judgments from photographs.

Tests involving miscellaneous abilities. Another of the excellent test approaches in the Horowitz's study (37) was labeled the *Categories Test*. This test was modeled directly upon a typical intelligence test item. Five pictures were presented, and the question asked "Which one does not belong?" As they adapted it, the item could be answered in more than one way, and the choice indicate something of the important social categories for the child. For example, a page might contain pictures of three white boys, one white girl, and one negro boy. In this instance, the child could use either sex or race as the dominant category. Other items provided opportunities to categorize by age and socio-economic status as well. Here again, the emphasis was on group norms rather than individual differences.

Hartley (32) has more recently used the categorization of photographs as a measure of ethnic salience. In this study, however, the task was presented in a voluntary framework, in contrast with the objective assignment presented in the *Categories Test*. In Hartley's "faces test" the respondent is told to: "Sort them out into piles. You can make as many piles as you want to, or as few as you choose. You can classify them on any basis you want to!"

In the Murray and Morgan "Study of Sentiments" (59) several of the "indirect methods, methods which conceal the examiner's intent" provide an ability test façade. The "Sentiments Examination" Section A Part 1 asks for the "most descriptive adjectives" for 48 stimulus words, "the S being led to believe that the examiner is interested in testing the range of his vocabulary." In Part 3 "The S is led to believe his verbal ability is being tested with a simile completion test," e.g. "As pathetic as a . . ." In the "Arguments Completion Test" the respondent is asked to continue and finish an argument, the beginning of which has been described. "Being led to believe his powers of argumentation are being tested, the subject quickly becomes involved in the controversy he or she is inventing, and ends by exposing more of his own sentiments than he might otherwise have done" (59, pp. 58-60).

Tests involving miscellaneous judgments. The following tests are disguised and structured, but in part at least must be classified as *voluntary*, in so far as judgments are required in situations in which there is no "objective" right answer. On the other hand, they retain the advantage of having the respondent (1) work on a task presumably less threatening than the experimenter's primary problem, and (2) report upon external values or realities, rather than upon himself directly. These studies like many already discussed, depend upon the demonstration of a judgment differential of which the respondent is presumably unaware.

In Watson's (81) "Moral Judgments" sub-test, judgments of approval or disapproval are made about a variety of situations, sets of these situations being identical except for the specific persons or groups involved. For example unwarranted search is made of a suspected "radical" headquarters, on the one hand, while in a parallel item, the same type of search is carried out with a business corporation suspected of dishonesty. Scoring is done on the basis of discrepancy of judgment between the parallel situations.

Similar in title and in plan to Watson's Moral Judgments Test is the ingenious approach to Negro-White attitudes devised by Seeman (72). While he used equated comparison groups, the technique could be modified for diagnos-

ing bias in individuals. He selected six items from a standard test of moral evaluations on marriage and sexual matters. To one group, the episode items were illustrated by pictures of white couples; to the other, with Negro illustrations. Both groups were made up of white college students. Contrary to expectation, the judgments were more lenient—less disapproval—with the Negro illustrations. Furthermore, when the two groups were sub-classified according to scores on the Likert scale of attitudes toward the Negro, the major part of the differential was contributed by the more *tolerant* persons rather than the more anti-Negro parts of the two groups. The intolerant extremes in this sample were more consistent, less "biased." These results are important and meaningful, but further indicate the danger of over-simplified interpretations in indirect approaches.

The readiness with which judgments of literary merit can be manipulated by the substitution of fictitious authors has been demonstrated by Farnsworth (70) and Sherif (73). Prestige seems to be one of those forces which can bias the performance of an objective task. Can we use this bias to infer attitudes? Lewis (47) seems to have had in mind some such approach. The present writer (22) deliberately set out to utilize the prestige effect to measure attitude toward five minority groups. Rather than literary passages, proverbs or mottoes were used. Rather than individual authors, the adage was attributed to the group as a whole, e.g., "American pioneer saying" or "old Jewish motto." Eight and ninth grade students were asked to evaluate the quality of each motto separately—there being ten mottoes attributed to each group. The test yielded a general prejudice score that correlated but .30 with a direct test. Other scores were worthless. However, in view of the strength of prestige suggestion in other experiments, this approach seems worth trying further. It is quite possible that in the form given the task was trivial and the disguise thin.

Horowitz (35) illustrated the development of patriotism and conformity among children by tests in which the task was to pick "the best looking flag," or the "best place to live in" from a series of photographs of houses, etc.

As Wolff, Smith and Murray (83) have shown, reactions to group-disparagement jokes are correlated with group membership. Gordon (29) has recently tried to utilize this phenomenon in the assessment of social attitudes. Twenty four jokes, both antagonistic and sympathetic, dealing with Negroes and Jews were rated as to their funniness on a 5-point scale. Four groups of college men were used: a Protestant fraternity, a Catholic fraternity, a Negro fraternity, a Jewish (non-Zionist) fraternity, and a Zionist club. The groups differed in their responses to these jokes but not as anticipated in all instances. The jokes sympathetic to Jews were rated highest by the two Jewish groups—but the anti-Jewish jokes failed to differentiate the groups. The anti-Negro jokes were rated lowest by the Negro group, but the pro-Negro jokes failed to differentiate. The various groups ranked the individual jokes quite similarly as to popularity, with rho's between the orders for the different groups ranging from .71 to .94. With regard to individual differences *within* the various groups, the ratings of the jokes showed no relationship to attitudes toward Negroes or Jews as revealed in a direct attitudes test or a symbol endorsement test. The tests were administered as a part of a larger testing program.

Cattell (13), in the program already referred to used a "Projection-Phantasy" test, in which the respondent picked from ten sets of ten incomplete sentences, five each which he completed. Rather than scoring the words used

in completing the sentences, Cattell scored the *choice* of incomplete sentences. Each set contained one representative of each of ten different attitudes. Applied to the desire to excel in one's profession, and to the desire for more sleep and rest, this test had Spearman-Brown reliabilities of .13 and .60. Correlations with the paired-comparison preferences were .15 and .31, with time and money expenditure, .01 and .26, with information, .20 and .20. In spite of the relatively low level of these values, this method was regarded as one of the more promising.

Kalpakian's "The Construction of a Disguised Test by the Use of Photographs for the Study of Attitudes Toward Negroes" (40) represents an attempt at disguise which fails in part to meet the criteria suggested here. Like other tests reported in this section, a judgment differential of which the respondent was unaware was sought. Twenty-six matched pairs of pictures were selected—one set depicting Negroes, the other whites. The test was structured, in that a choice of responses, "Like—Indifferent—Dislike" was required for each picture. The instructions read in part "we're interested in seeing what kinds of pictures people like and don't like . . . we're not interested in the technical photographic aspects of the pictures, rather we want your reactions to what is in the pictures. That is, how do you feel about what is in each picture." Note the voluntary nature of this assignment—the respondent being told in effect that there are no right or wrong answers. Note also the absence of any clear cut assigned basis of judgment which might lend indirection (such as asking for judgments of beauty of the photograph or skill of photographer). Kalpakian did aid the disguise by placing these 52 photographs in random order among 68 irrelevant pictures. This is similar to the efforts in disguise made by those who distribute the items of one attitude scale among a variety of other types (e.g., 24). In one of Kalpakian's groups, 100 subjects were asked what they thought the purpose of the test was. Only 15 did not guess that the purpose involved attitudes toward the Negro; 51 thought this and several other attitudes were being tested; 34 judged correctly that the primary purpose was attitudes toward Negroes. On the other hand only 7 were aware of the matching of the pictures. Scored on proportion of white to Negro "likes" the test showed corrected split-half reliability coefficients between .71 and .82 in various white high school and college groups. Mean values showed a definite "white salience" for all groups. Scores correlated with both self ratings and ratings by friends, demonstrated by critical-ratio tests. The correlation with the Thurstone-Hinkley was .40, which became .80 when corrected for restricted range. In terms of the categories of this review, Kalpakian's test is similar to the "show me" and other picture tests developed by Horowitz (34). It is structured in that a limited number of response categories is provided. It is direct, substituting pictures for the ethnic labels used in statement-endorsement type of direct tests.

GENERAL CONSIDERATIONS

Validity of indirect tests. Most efforts to develop indirect attitude tests are predicated upon the assumption that indirect tests will under certain conditions have higher validity than direct tests. It is worthy of note that none of the studies reviewed here offers any evidence that

this is the case. The testing of this hypothesis is certainly the next step if these preliminary efforts are thought to be sufficiently promising to justify more research.

The most usual effort at validation has involved correlating the indirect test with direct test results, under conditions in which the latter were gathered anonymously or in which there was no reason to suspect respondent dissembling (e.g., 19, 22, 29, 38, 40, 43, 60, 64, 65). Hammond (31) and Gordon (29) used environmentally selected criterion or comparison groups. Parrish's tests (64) meaningfully reflected an experience differential between two groups. These are, however, isolated instances, and even at the exploratory level, we might expect more use of criterion groups.

As has been pointed out previously, the case for *face validity* for indirect tests may often be better than for direct tests. In a number of the disguised, structured tests, the distribution of scores and measures of internal consistency demonstrates unequivocally that non-random, systematic errors, differences in perceptions, etc. exist, of which the respondents are presumably unaware. These systematic unconscious "biases" are well worth study in their own right, and seem to lie close to the functional meaning of attitude.

Attitude measurement as related to attitude definition and theory. Research on social attitudes has been justly criticized for a lack of common definition of the concept, and for a failure to integrate definition and measurement procedures. This diversity of definition has been in odd contrast with the obvious similarity of research procedures. This paradox arises from definitional attempts which confound *explanations* of the phenomena with the process of *pointing* to the phenomena. It is the contention of the present writer that agreement on the implicit operational (or pointing) definition of attitudes is already present. As a tentative formulation the following is offered: *A social attitude is (or is evidenced by) consistency in response to social objects.* If we look at those definitions utilizing concepts of set, or readiness to respond—for example, Allport's (1): "An attitude is a mental and neural state of readiness, organized through experience and exerting a directive or dynamic influence upon the individual's response to all objects and situations to which it is related"—and ask for the evidence of a "mental and neural state of readiness," the symptoms of a "directive or dynamic influence," criteria as to the "objects and situations to which it is related," these evidences will be in final analysis consistency or predictability among responses. *An individual's social attitude is a syndrome of response consistency with regard to social objects.* And even those whose be-

havioristic orientation leads to a rejection of such mentalistic definitions as Allport's—and who would say with Bain (3) and Horowitz (36, p. 142), "essentially . . . the attitude must be considered a response rather than a set to respond"—in research practice do not equate *isolated* responses with attitudes, but on the contrary look for the appearance of *response consistencies*. This is dramatically evidenced by Horowitz's (34) use of the appearance of consistent differentiated response to photographs of Negro and white children to mark the occurrence of race prejudice in children.

This standpoint provides a basis for the operational delimitation of specific and general attitudes. It likewise demands that if attitude measurement be integrated with definition, internal consistency among the sample responses collected by the test must be demonstrated. To the present writer, the demonstration of a single factor through factor analysis seems ideal, although time consuming. Spearman-Brown and Kuder-Richardson (44) reliability measures, Guttman's (30) "reproducibility coefficient," Loevinger's (52) "homogeneity," and Thurstone's "coefficient of irrelevance" (78) are likewise evidences of internal consistency. Item analysis against total scores (49, 75), remains one of the best practical means of improving internal consistency.

Contemporary American theories commonly hypothesize attitudes to be learned, or identify them with other cognitive phenomena such as perception and concept formation. From the operational definition suggested it would be hard to distinguish attitude from habit, which likewise is evidenced by behavior consistency. Indeed, if learning be conceived broadly enough to include the development of any response consistency, attitude formation is identical with learning. Traditionally, learning has been used in those situations in which the response consistencies represented optimal reflections of a simple, uniform environment—where increasing response consistency represented the elimination of "error." But a similar process goes on in "unsolvable" or highly complicated environments, and stereotyped modes of response are developed which are non-optimal reflections of the environment, which do not represent error-reduction. The terms perseveration, set, fixation, position habit, bias, irrational behavior, idiosyncrasy, pre-judgment, non-random error, over simplification, bad habit, rigidity, all come to mind to describe this process. Presumably, this latter type of response consistency is a result of the same cognitive processes which under other conditions result in "right answer" learning. It is with these common processes that the concept of attitude is identified, although typically it is most closely associated with the acquisition of "irrational" or non-adaptive response consistencies.

In these terms it may be meaningful to compare direct and indirect efforts at attitude assessment. In both direct and indirect tests enough internal consistency has been demonstrated on some topics to justify speaking of attitudes. Likewise, in numerous instances reported here, considerable correlation has been demonstrated between responses to direct and to indirect tests, justifying the hypothesis that a common attitude lies behind them. Missing is the demonstration that this pattern of consistency extends beyond paper and pencil to the "real-life" situations which are usually in mind when the concept of "attitude" is used. While meaningful research problems exist in regard to patterns of consistency within a universe of paper and pencil responses, the bulk of the researches reviewed are predicated upon the assumption that broader consistencies exist—that overt behavioral manifestations of the attitude can be predicted.

With regard to the level of response consistency, it should be noted that direct tests uniformly have much higher reliability coefficients than do indirect ones, especially when the number of items and time of administration are considered. Of course this consistency is in part conscious, voluntary, and possibly superficial—in contrast to the involuntary "bias" in performance achieved by many indirect tests. In this sense, the indirect tests of the non-voluntary sort do in themselves demonstrate and utilize the dramatic phenomena of attitudinal interference with tasks of learning, perceiving, remembering, and evaluating, such as are summarized by Sherif and Cantril (74, pp. 29-91). And it is such consistencies, rather than consistency in voluntary self-description, which give vitality to current usage of the concept of attitude. In this, indirect attitude tests lie closer to contemporary attitude theory than do direct tests.

SUMMARY

The efforts at indirectly measuring social attitudes have been summarized. A distinction has been offered between disguised, non-structured, types of test on the one hand, and disguised, structured, tests on the other. Because of a greater freedom from dependence upon the respondent's voluntary self description, the use of disguised, structured tests has been emphasized.

As can be seen from the survey, there is an abundance of partially tried techniques. Equally good ones can be developed from the general formula: A plausible task, (a) which your respondents will all strive to do well, (b) which is sufficiently difficult or ambiguous to allow individual differences in response, and (c) which can be loaded with content relative to the attitude you seek to measure. Test the responses of individuals for persistent selectivity in performance, for correlated

or non-random errors. For example, several persons have suggested to the present writer what might be called the "photocrime" test. Here criminal scenes would be pictured in photograph and text. Sufficient ambiguity would be provided to make it a challenging problem to the amateur detective, and, incidentally, to allow the possibility of bias in the selection of the criminal from the suspects of various ethnic or social identification. A test of "common legal knowledge" might well provide a framework for carrying further some of the notions in Watson's "Moral Judgments" sub-test. The respondent might be asked to decide whether or not civil liberties had been violated in a wide range of specific situations—might be asked to decide between claimant and defendant in a number of civil suits. Here the content could easily be manipulated to get at almost any social attitude one might be interested in. It should not be too difficult to keep the respondent convinced that he dealt solely with matters of the law. We are prone to see as propaganda that with which we disagree. A "propaganda analysis test" should reveal indirectly some people's attitudes. For the build-up, give a review of some principles on how to spot propaganda, how to detect an ulterior motive, a selfish interest group, deceit, exaggeration, etc. Then have the respondent test his skill on a number of argumentative paragraphs, which he evaluates as to the extent and kind of propaganda present. The root of this notion may be found in Watson's "Arguments" sub-test.

Such devices may be multiplied; the reader can no doubt provide better ones. In some of these, the fiction of an objective test situation will be hollow indeed. In others, for example in the estimation of group opinion, we are tapping biased performance in an everyday act of social perception. We are not taking at face value the respondent's own volunteered description of himself. Rather we are studying systematic errors in the respondent's own perceptions, errors of which he himself is not aware. Not only in attitude testing but in personality study in general, distortions of performance in dealing with the environment provide objective evidence of an individual's unique picture of his world. Just as the psychophysicists were long ago forced to introduce *Vexirversuche*, or catch trials, into introspection—to eliminate impossible negative reaction times and infinitesimal two-point thresholds—so personality testing, whether check-list or free-response, must rise above dependence on voluntary self-description. The development of structured indirect attitude tests is thought to be a step in this direction.

BIBLIOGRAPHY

1. ALLPORT, G. W. Attitudes. Ch. 17 in C. Murchison (Ed.), *A handbook of social psychology*. Worcester, Mass.: Clark University Press, 1935.
2. ALLPORT, G. W., & POSTMAN, L. *The psychology of rumor*. New York: Holt, 1947.
3. BAIN, R. An attitude on attitude research. *Amer. J. Sociol.*, 1928, 33, 940-957.
4. BARTLETT, F. C. *The study of society*. London: Kegan Paul, 1939.
5. BARTLETT, F. C. The cooperation of social groups: a preliminary report and suggestions. *Occup. Psychol.*, 1938, 12, 30-42.
6. BROWN, J. F. Modification of the Rosenzweig picture frustration test to study hostile interracial attitudes. *J. Psychol.*, 1947, 24, 247-272.
7. BURTT, H. E. Measuring interests objectively. *Sch. & Soc.*, 1923, 17, 444-448.
8. BURTT, H. E. *Principles of employment psychology* (Rev. Ed.). New York: Harper, 1942.
9. CARMICHAEL, D. M. The cooperation of social groups. *Brit. J. Psychol.*, 1938, 29, 206-231; 329-344.
10. CATTELL, R. B. *A guide to mental testing*. London: Univ. London Press, 1936.
11. CATTELL, R. B. Projection and the design of projective tests of personality. *Charact. and Pers.*, 1944, 12, 177-194.
12. CATTELL, R. B., HEIST, A. B., HEIST, P. A., & STEWART, R. G. The objective measurement of dynamic traits. (To be published in *Educ. psychol. msmt.*, 1950.)
13. CATTELL, R. B., MAXWELL, E. F., LIGHT, B. H., & UNGER, M. P. The objective measurement of attitudes. (To be published in the *Brit. J. Psychol.*, 50, 1950.)
14. COFFIN, T. E. Some conditions of suggestion and suggestibility: a study of certain attitudinal and situational factors influencing the process of suggestion. *Psychol. Monogr.* 1941, 4, No. 241.
15. COLLEGE STUDY IN INTERGROUP RELATIONS. (Lloyd Allen Cook, Director.) Study forms and technics in intergroup relations. Supplementary Sheet No. 5, Jan., 1948. (Mimeographed.) Detroit: Wayne University.
16. COMMISSION ON COMMUNITY INTERRELATIONS OF THE AMERICAN JEWISH CONGRESS. (Stuart Cook, Research Director.) The face game. Unpublished research on a modified version of a test originally developed by Marian Radke. 1947-1948.
17. DAVIS, T. E. Some racial attitudes of Negro college and grade school students. *J. Negro Educ.*, 1937, 6, 157-165.
18. DERI, S., DINNERSTEIN, D., HARDING, J., & PEPITONE, A. D. Techniques for the diagnosis and measurement of intergroup attitudes and behavior. *Psychol. Bull.*, 1948, 45, 248-271.
19. DUBIN, S. S. Verbal attitude scores predicted from responses in a projective technique. *Sociometry*, 1940, 3, 24-28.
20. EDWARDS, A. L. Political frames of references as a factor influencing recognition. *J. abnorm. soc. Psychol.*, 1948, 36, 34-61.
21. EVANS, M. C., & CHEIN, I. The movie story game: a projective test of interracial attitudes for use with Negro and white children. Paper read at the 56th annual meeting of the American Psychological Association, Boston, Sept. 8, 1948.
22. FRENKEL-BRUNSWIK, E., JONES, H. E. (Directors), & ROKEACH, M., JARVIK, M., & CAMPBELL, D. T. (Staff). Unpublished research on the personality correlates of anti-mi-

- nority attitudes among grade school children. A project of the University of California Institute of Child Welfare, financed by a grant from the American Jewish Committee, 1946-1947.
23. FRENKEL-BRUNSWIK, E., & REICHERT, S. Personality and prejudice in women. (Unpublished manuscript.)
 24. FRENKEL-BRUNSWIK, E., LEVINSON, D., & SANFORD, R. N. The anti-democratic personality. In T. M. Newcomb & E. L. Hartley (Eds.), *Readings in social psychology*. New York: Holt, 1947.
 25. FROMME, A. On use of qualitative methods of attitude research. *J. soc. Psychol.*, 1941, 13, 429-460.
 26. FRYER, D. *The measurement of interests*. New York: Holt, 1931.
 27. GILBERT, H. H. Secondary science and pupil prejudice. *J. educ. res.*, 1941, 35, 294-299
 28. GLAZER, E. M. An experiment in the development of critical thinking. *Teach. Coll. Contr. Educ.*, No. 843. New York: Bureau of Publications, Teachers Coll., Columbia Univ., 1941.
 29. GORDON, S. Exploration of social attitudes through humor. Master's Thesis, Univ. of Illinois, 1947.
 30. GUTTMAN, L. A basis for scaling qualitative data. *Amer. sociol. Rev.*, 1944, 9, 139-150.
 31. HAMMOND, K. R. Measuring attitudes by error-choice; an indirect method. *J. abnorm. soc. Psychol.*, 1948, 43, 38-48.
 32. HARTLEY, E. L. *Problems in prejudice*. New York: Kings Crown Press, 1946.
 33. HARTLEY, E. L., & SCHWARTZ, S. A pictorial-doll play approach for the study of children's intergroup attitudes. Mimeographed preliminary draft. Research Institute in American Jewish Education, American Jewish Committee, Summer, 1948.
 34. HOROWITZ, E. L. The development of attitude toward the Negro. *Arch. Psychol.*, N. Y., 1936, 28, No. 194.
 35. HOROWITZ, E. L. Some aspects of the development of patriotism in children. *Sociometry*, 1940, 3, 329-341.
 36. HOROWITZ, E. L. 'Race' attitudes. In O. Klineberg (Ed.), *Characteristics of the American Negro*. New York: Harper, 1944.
 37. HOROWITZ, E. L., & HOROWITZ, RUTH E. Development of social attitudes in children. *Sociometry*, 1938, 1, 301-338.
 38. HSÜ, E. H. An experimental study of rationalization. *J. abnorm. soc. Psychol.*, 1949, 44, 277-278.
 39. JOHNSON, G. G. An experimental analysis of the origin and development of racial attitudes with special emphasis on the role of bilingualism. Ph.D. Thesis, Univ. of Colorado, 1949.
 40. KALPAKIAN, E. Y. The construction of a disguised test by use of photographs for the study of attitudes toward Negroes. Master's Thesis, Clark University, 1947. (*Clark University Bulletin*. Abstracts of Dissertations and Theses, Worcester, Mass., Oct., 1947.)
 41. KATZ, MARTIN R. A hypothesis on anti-Negro prejudice. *Amer. J. Sociol.*, 1947, 53, 100-104.
 42. KRECH, D., & CRUTCHFIELD, R. S. *Theory and problems of social psychology*, New York: McGraw-Hill, 1948.
 43. KREMEN, E. O. An attempt to ameliorate hostility toward the Negro through role playing. Master's Thesis, The Ohio State University, 1949.
 44. KUDER, G. F., & RICHARDSON, M. W. The theory of the estimation of test reliability. *Psychometrika*, 1937, 2, 151-160.
 45. LANDSDELL, H. A study of distorted syllogistic reasoning as a means of discovering covert attitudes toward marriage. *Bull. Canad. Psychol.*

- Assn.*, 1946, 6, 98. (Abstract.)
46. LEFFORD, A. The influence of emotional subject matter on logical reasoning. *J. gen. Psychol.*, 1946, 34, 127-151.
 47. LEWIS, H. B. An approach to attitude measurement. *Psychol. League J.*, 1938, 2, 64-67.
 48. LEVINE, J. M., & MURPHY, G. The learning and forgetting of controversial material. *J. abnorm. soc. Psychol.*, 1943, 38, 507-517.
 49. LIKERT, R. A technique for the measurement of attitudes. *Arch. Psychol.*, N. Y., No. 140, 1932.
 50. LOEBLOWITZ-LENNARD, H., & RIESSMAN, F., JR. A proposed projective attitude test. *Psychiatry*, 1946, 9, 67-68.
 51. LOEBLOWITZ-LENNARD, H., & RIESSMAN, F., JR. A preliminary report on social perception test—a new approach to attitude research. *Social Forces*, 1946, 24, 423-427.
 52. LOEVINGER, J. The technic of homogeneous tests compared with some aspects of "scale analysis" and factor analysis. *Psychol. Bull.*, 1948, 45, 507-529.
 53. MCGREGOR, D. The major determinants of the prediction of social events. *J. abnorm. soc. Psychol.*, 1938, 33, 179-204.
 54. McNEMAR, Q. Opinion-attitude methodology. *Psychol. Bull.*, 1946, 43, 289-374.
 55. MORGAN, J. J. B. Distorted reasoning as an index of public opinion. *Sch. & Soc.*, 1943, 57, 333-335.
 56. MORGAN, J. J. B., & MORTON, J. T. The distortion of syllogistic reasoning produced by personal convictions. *J. soc. Psychol.*, 1944, 20, 39-59.
 57. MORGAN, J. J. B. Attitudes of students toward the Japanese. *J. soc. Psychol.*, 1945, 21, 219-246.
 58. MURRAY, H. A. *Explorations in personality*. New York: Oxford Univ. Press, 1938.
 59. MURRAY, H. A., & MORGAN, C. D. A clinical study of sentiments. I and II. *Genet. Psychol. Monogr.*, 1945, 32, 3-311.
 60. MURPHY, G., & LIKERT, R. *Public opinion and the individual*. New York: Harpers, 1937.
 61. NEWCOMB, T. M. *Personality and social change*. New York: Dryden Press, 1943.
 62. NEWCOMB, T. M. The influence of attitude climate upon some determinants of information. *J. abnorm. soc. Psychol.*, 1946, 41, 291-302.
 63. NOLAND, E. W. Factors associated with absenteeism in a south central New York State industry. Ph.D. thesis on file in the Cornell University Library, 1944.
 64. PARRISH, J. A. The direct and indirect assessment of attitudes as influenced by propagandized radio transcriptions. Master's thesis, The Ohio State University, 1948.
 65. PROSHANSKY, H. A projective method for the study of attitudes. *J. abnorm. soc. Psychol.*, 1943, 38, 383-395.
 66. ROBBINS, I. Point of view and quality of thought in attitude measurement. *Improving Educational Research*, pp. 52-56. American Educ. Res. Ass. 1948 Official Report, Washington, D. C.
 67. ROSE, ARNOLD. *Studies in reduction of prejudice*. Chicago: American Council on Race Relations, 1948.
 68. ROTTER, J. B., & WILLERMAN, B. The incomplete sentences tests as a method of studying personality. *J. consult. Psychol.*, 1947, 11, 43-48.
 69. RUCH, F. L. *Psychology and life*. (1st Ed.). Chicago: Scott, Foresman and Company, 1937.
 70. SAADI, M., & FARNSWORTH, P. R. The degrees of acceptance of dogmatic statements and preferences for their

- supposed makers. *J. abnorm. soc. Psychol.*, 1934, **29**, 143-150.
71. SEELEMAN, V. The influence of attitude upon the remembering of pictorial material. *Arch. Psychol.*, N. Y., 1940-1941, **36**, No. 258.
 72. SEEMAN, M. Moral judgments: a study in racial frames of references, *Amer. sociol. Rev.*, 1947, **12**, 404-411.
 73. SHERIF, M. A study of some social factors in perception. *Arch. Psychol.*, N. Y., 1935, No. 187.
 74. SHERIF, M., & CANTRIL, H. *The psychology of ego-involvements*. New York: Wiley, 1947.
 75. SLETT, R. F. *Construction of personality scales by the criterion of internal consistency*. Hanover: The Sociological Press, 1937.
 76. SMITH, G. H. Beliefs in statements labeled fact and rumor. *J. abnorm. soc. Psychol.*, 1947, **42**, 80-90.
 77. SWEET, L. *The measurement of personal attitudes in younger boys*. New York: The Association Press, 1929.
 78. THURSTONE, L. L., & CHAVE, E. J. *The measurement of attitude*. Chicago: Univ. Chicago Press, 1929.
 79. TRAVERS, R. M. W. A study in judging the opinions of groups. *Arch. Psychol.*, N. Y., No. 266, 1941.
 80. WALLEN, R. Individuals' estimates of group opinion. *J. soc. Psychol.*, 1943, **17**, 269-274.
 81. WATSON, G. B. The measurement of fair-mindedness. *Teach. Coll. Contr. Educ.* No. 176. New York: Teachers Coll., Columbia Univ., 1925.
 82. WILLIAMS, ROBIN M., JR. *The reduction of intergroup tensions*. New York: Social Science Research Council Bull., No. 57, 1947.
 83. WOLFF, H. A., SMITH, C. E., & MURRAY, H. A. The psychology of humor; 1. a study of race disparagement jokes. *J. abnorm. soc. Psychol.*, 1934, **28**, 341-365.
 84. ZELIGS, R. Racial attitudes of children. *Sociol. soc. Res.*, 1937, **21**, 361-371.

Received May 1, 1949.

AN EXPERIMENTAL ANALOGUE OF REPRESSION

I. HISTORICAL SUMMARY¹

ANCHARD FREDERIC ZELLER

The Johns Hopkins University

INTRODUCTION

At the heart of the theoretical psychoanalytic structure is the concept of repression, which is the mechanism by which painful or unpleasant material is excluded from consciousness and motor expression. In his earlier writings Freud (17) stressed the importance of repression as the primary ego defense but in his later works he revised his concept by pointing out that repression is only one of several mechanisms which the ego may utilize in an effort to avoid unpleasantness. The ego may indulge in flight, it may resort to condemnation, in which case the material remains conscious, or it may resort to repression.

Repression is of two kinds: first, primal or archaic repression which denies entrance into consciousness of some archaic idea attached to instinctual strivings which are unacceptable to the ego, and second, after-expulsion or repression proper which pushes from consciousness material which, although once conscious, has in some way become associated with primally repressed material. Repression is not forgetting. It is an active process which requires the exertion of constant energy by the ego. This is true for both primal repression and after-expulsion. Repression may be complete or partial. If it is complete there is very little hope of investigating it experimentally. If it is partial the repressed material may remain as it was, ready to reappear in consciousness if the repression is removed, or the force attached to it may manifest itself in consciousness in the form of neurotic symptoms. Incomplete repression lends itself more readily to experimental investigation. From the standpoint of the clinician, after-expulsion is of more importance than primal repression, since most observed cases of repression are examples of incomplete after-expulsion. That is, an individual is unable to recall material which he has at one time known but which has in some way become unacceptable, but the memory for which may be restored by appropriate treatment. Although seldom explicitly stated, this is

¹ The present paper is a portion of a dissertation presented to the Faculty of Philosophy of the School of Higher Studies, The Johns Hopkins University, in partial fulfillment of the requirements for the Ph.D. degree. The writer is indebted to Professors Stanley B. Williams and W. C. H. Prentice for advice and assistance during the course of the study.

the concept of repression which has served as a basis for most of the experiments in the field. A great many experiments have been designed, first, to demonstrate the phenomenon and second, to quantify the factors involved. Such a demonstration must of necessity fulfill three requirements. First, it must demonstrate that the material in question has been learned by the individual. Second, it must demonstrate that the introduction of an inhibiting factor causes inability to recall or a significant decrease in the recall of the material. Third, it must show that the removal of the inhibiting factor results in the reinstatement of the ability to recall the material. As will be shown in this review of the literature there has as yet been no clear cut laboratory demonstration which includes all of these steps.

EXPERIMENTAL STUDIES OF REPRESSION

Questionnaire method. The first experimental attack was made by F. W. Colgrave (10) in 1898 when he administered a questionnaire to a group of school children which contained the question, "Do you recall pleasant or unpleasant experiences better?" He concluded from his results that pleasant items were better recalled than unpleasant. As critics have pointed out, the questionnaire method is not a highly valid measure, but a start had been made. Kowalewski (30) and again Susukita (81) as late as 1935 also utilized the questionnaire in the study of this problem.

Associations with sensory stimuli. In 1905 Gordon (20), working in Külpe's laboratory, initiated another type of attack on the problem by associating sensory stimuli with material to be learned and later testing recall. She used colored figures paired with protocols, and administered a recognition test after three weeks. She found no difference between pleasant, unpleasant, and indifferent items. Tait (82), using color recognition, found a slight superiority for memory associated with pleasant colors. Anderson and Bolton (1) tested the recognition and recall of odors with nonsense names attached and found little difference between pleasant, unpleasant, and indifferent items. Gordon (21) again attacked the problem in 1925, this time using pleasant and unpleasant odors followed by a memory test. Again she found no difference between memory for pleasant and unpleasant items. Kenneth (28) used odors as a free association stimulus. He found more pleasant than unpleasant associations, but the results are not especially relevant to repression. Frank and Ludvigh (16) paired odors with nonsense syllables and found that pleasant associations were recalled oftener. The most complete experiment on memory value of pleasant and unpleasant items was conducted by Ratliff (55) using sensory stimulation paired with numbers. Three types of sensory stimuli, odors, colors, and consonant chords, were paired with the

numbers. He found that with visual and auditory stimulation the pleasant was recalled more often than the unpleasant, but for olfaction the unpleasant recall was superior.

Recall of experiences. Another approach extensively used to study the relation of affect to memory was initiated by Kowalewski (31) in 1908 when, the day following a Christmas vacation, he instructed his students, "Write down whatever pleased you yesterday." His results indicated more pleasant than unpleasant experiences. A recall ten days later gave similar results. He interpreted his results to mean that the pleasant is better retained than the unpleasant. As critics were quick to point out, however, he assumed that pleasant and unpleasant experiences were equal in number, which assumption has more recently been proven erroneous. Other investigators employing the recall of holiday experiences have been Wohlgemuth (91), Meltzer (40, 41, 43), Jersild (27), Cason (8), Waters and Leeper (87), Menzies (44), O'Kelly and Steckle (48), and Steckle (79). The studies of Wohlgemuth (91) and Cason (8), two of the best controlled, indicate no difference between recall of pleasant and unpleasant associations when the fact that there are actually more pleasant than unpleasant experiences is taken into account, while the results of the other investigations indicate only a slight tendency for greater recall of pleasant experiences.

Kowalewski (30, 31) concluded from his results that the recall of affective material is related to personality types. He interpreted his data to mean that optimists remember proportionally more pleasant experiences while pessimists tend to remember proportionally more unpleasant experiences. This possibility has been investigated by Meltzer (40, 43), O'Kelly and Steckle (48), and Steckle (79) with inconclusive results. Flugel in 1917 (see Flugel, 1925, 14) attempted to establish an "Algedonic Ratio" of pleasant to unpleasant for the individual. This was again attempted by Meltzer (41) in 1930.

Other experiments which have used the recall of past experience have been conducted by Gordon (22) and Thompson (83) who used the recall of childhood experiences. Gordon found no evidence for a greater percentage of pleasant recall, but Thompson found evidence for a pleasant-unpleasant differential in favor of the pleasant.

Discrete associations. In addition to the experiments in which preceding events have been recalled and re-recalled, a number of experiments have utilized other methods to test memory for affectively toned experience. Peters (49) had subjects answer to a stimulus word with an experience. He found 65 per cent of total recalls were pleasant in nature. Peters and Nemecek (50) used the same technique with comparable results. Griffitts (24) asked for responses to stimulus words and found a difference in favor of pleasant associations. Stagner (77) in a well-controlled experiment in which a single pleasant and unpleasant event were recorded, together with their associations, found what he called

evidence for an active repression when the events were again presented to the subjects with instructions to recall the associations.

Reaction time. The effect on reaction time of various affectively toned material has been investigated by Birnbaum (5), Tolman (84), Tolman and Johnson (85), Baxter, Yamada and Washburn (4), Morgan, Mull and Washburn (46), and Smith (76). All but Birnbaum found positive results, viz., pleasant tone decreases reaction time.

Memory for word lists. The number of experiments using this method has been large. Chaney and Lauer (9), Lynch (36), Cason (8), Stagner (78), Balken (2), Bunch and Wientge (6), Silverman and Cason (75), White and Ratliff (90), Carter (7), White (89), Gilbert (18), Lanier (33), and Pintner and Forlano (51) have all worked with this type of material, using periods of recall up to two years. None of these experiments has given a conclusive answer to the problem. They have all tended to show, however, that the problem of affect and recall is not a simple one, but rather a very complex phenomenon depending on many factors, such as sex, age, social status, intelligence, etc. (see Gilbert, 1938).

Miscellaneous studies. A few other studies have been conducted which do not readily fall into the above categories. Laird (32) asked his subjects to write the names of ten friends and check those liked and disliked. Later he asked the subjects to list ten names as fast as possible. A comparison of the two lists indicated that some subjects tended to list more disliked persons in the second list than did others. He interpreted his results in terms of optimistic and pessimistic personality types but found no evidence for repression. Fox (15) had his subjects learn pleasant and unpleasant sonnets and found those preferred were better retained. Stone (80) used photographs to which names liked or disliked by the subjects were attached. He found that those associated with pleasant names were better recalled. McMullin (39) found evidence that interpolated activity of an affective nature could produce differences in the amount of retroactive inhibition.

Review articles or discussions of various aspects of pleasant-unpleasant experiments have been presented by Meltzer (42), Moore (45), Sears (64, 67, 68, 69), Gilbert (19), McGeoch (37), Rapaport (54), and Prentice (53). As can be seen from the experiments reviewed so far, very little has been contributed to an understanding of the problem of repression as outlined by Freud. Of the experiments mentioned so far, the score stands 32 to 14 in favor of more effective recall of pleasant than of unpleasant experiences, with 5 experiments having neutral results. These results, however, can not be taken as an indication of repression.

As Meltzer (42) has pointed out, most of the experiments before 1930 were conducted on the false assumption that the pleasant and unpleasant events were of equal frequency. Henderson (25) as early as

1911, however, recognized the fallacy and interpreted this as failing to support the idea that the unpleasant was repressed. Myers (47) also argued that we do not forget the disagreeable, but that there are more agreeable than disagreeable experiences in life. Flugel (14) also recognized this as did Griffiths (24) and Wohlgemuth (91). In none of the studies has any attempt been made to restore recall when repression has been assumed as basic to the pleasant-unpleasant differential. In most cases repression has been equated to differential forgetting, whereas it could have been more reasonably interpreted as differential learning. In most cases repression has not been conceived as the dynamic process outlined by Freud, and in no case has an adequate test been designed.

Frame of reference. There have, however, been a number of studies which have recognized the importance of learning set. Following the work of Bartlett (3) a number of studies have been conducted in which controversial material about which the subjects were known to have specific opinions was presented and then recalled later. As early as 1928 Zillig (93) gave both men and women a number of selections to read, the content of some being favorable and others being unfavorable to women. The women recalled a much greater percentage of favorable items about themselves than the men did about the women. Studies following the same plan have been conducted by Watson and Hartmann (88) using theistic and atheistic material, Seeleman (70) using attitude toward the negro, Edwards (11) with pro- and anti- New Deal material, Levine and Murphy (34), pro- and anti-communistic literature and Postman and Murphy (52) using United Nations vs. Axis power attitudes. Wallen (86) used his subjects' own opinion of themselves, compared with bogus group opinions, and found they remembered the items which agreed with their own opinions better than those which did not. Shaw (73) and Shaw and Spooner (74) repeated the same experiment with modifications and obtained comparable results. All these studies indicate that attitude and preconceptions influence memory. Edwards (12) has presented an argument in favor of a "frame of reference" determinant in forgetting, regardless of whether the material is intrinsically pleasurable or otherwise. Steckle (79) presents a similar argument. Again, however, it should be noted that the analysis is of differential forgetting, not of an active, removable repression.

Specific studies. There have been a number of experiments based on different approaches to the problem which have shown considerably more insight into the nature of the problem and more imagination in the design of the experiment. Koch (29) using a number of psychology students had them rate their satisfaction with grades on quizzes. Five weeks later a recall test of the grades indicated the grades with the most affect, either positive or negative, were better recalled. In 1936 Sears (64) presented an excellent review of functional abnormalities

of memory in which he pointed out that none of the experiments had fulfilled the conditions of a true test of repression, since the fundamental assumption of the experiments had been that pleasantness and unpleasantness of an intellectual or sensory nature is equivalent to unpleasantness in terms of ego threat. In 1937 Sears (66) presented a study in which subjects were given two tasks, learning nonsense syllables and sorting cards. A list of syllables was learned, followed by success or failure at card sorting, and this was in turn followed by the learning of a second list. The second list learned by the successful card sorters was significantly better than the learning of the same list by those who failed at card sorting. Sears interprets this as evidence of repression. It might be pointed out that the difference could well be attributed to lesser motivation rather than to an active repression. No attempt was made to remove the repression.

One of the better demonstrations of repression has been given by Huston, Shakow, and Erickson (26) using the technique developed by Luria (35) in which the "associations of higher central nervous system processes are associated with voluntary movement so that the conflicts in the former are disclosed in the latter." The subjects were hypnotized and were told that they had participated in some event in a manner out of keeping with their normal standard of ethics. Stimulus words, some neutral and some related to the suggested experience, were read to the subjects who had a post-hypnotic amnesia for the suggestions. There were significant differences in reaction to the words associated with the hypnotic suggestions. The subjects were then rehypnotized and the suggestion removed. Retests indicated the effects had disappeared. This experiment comes close to a demonstration of repression, but the use of hypnosis as a repression medium makes it difficult to interpret.

Flanagan (13) and Sharp (71) at the University of Chicago used sexual and profane word pairs, respectively, and found significant differences in the learning and recall of the two types of material. It should be pointed out in this connection that no attempt was made to equate the affective and control material for difficulty in learning, and the experimental situation was such that embarrassment over the nature of the material to be reproduced undoubtedly contributed a great deal to the results. Sharp (72) in 1938, using neurotic subjects and material taken from case histories, found curves of retention which would indicate that pleasant material is more readily recalled after a time while unpleasant material is repressed (forgotten).

McGranahan (38) in an interesting interpretation of the Freudian concept, points out that repression is a process which can be carried out more efficiently at a conscious level by a well-integrated personality. To demonstrate this, he forbade his subjects to name a color on penalty of being shocked. The subject then answered to a series of stimulus words to which color responses are common. The well-integrated in-

dividuals "repress" the color responses while those who fear the shock make more color responses.

In 1930 Meltzer (42) concluded after a critical review of the studies on feeling and memory that the difficulty with repression studies was that the investigators were looking for universal responses. Rosenzweig (56, 57, 58), and Rosenzweig and Mason (61) pointed out that the lack of acceptable results in previous studies was due to a lack of clear understanding of the problem. They further point out the fallacy of assuming that sensory pleasantness and unpleasantness is the same as conative pleasantness and unpleasantness. Further, repression operates on a conative type of material which is negative in hedonic tone, the negative affect being determined by the conflict of the material with ego-supporting drives, such as self-respect, etc. They cite the experiments of Zeigarnik (92) as a prototype of the repression experiment and set up a series of investigations in which children were allowed to complete one half of the tasks and fail on one half. Recall was better for the completed tasks. Varying the task set, Rosenzweig (59, 60) found that ego-involvement led to recall of completed tasks, while task orientation led to better recall of uncompleted tasks. Two experiments conducted by Sarason and Rosenzweig (62, 63) in connection with Rosenzweig's triadic hypothesis further attempted to correlate repression with hypnotizability and impulsive reactions to conflict. Rosenzweig interprets his evidence as lending support to his hypothesis.

Gould (23) designed an experiment which came close to a test of repression. The subject was required to choose one of two tasks. The choice supposedly revealed a good or bad character trait of the subject, the character traits "revealed" having been predetermined by the investigator regardless of the task chosen. The subject was then told what the choice revealed about him and was allowed to complete the task, then make another choice until six had been chosen. The subject was then asked to list all of the tasks. Following this, the nature of the experiment was explained to the subject. As an incidental remark, Gould notes that after the explanation some of the subjects remembered more of the tasks than before.

SUMMARY AND CONCLUSIONS

A review of the literature shows that there has been no experiment which has fulfilled the criterion of a laboratory test of repression. Most experiments have been directed toward testing the relationship between affect and recall. The majority of experiments have been found to fall in eight categories:

1. Questionnaire method: this was the earliest used method and is unsuited to the problem.
2. Associations with sensory stimuli: the chief criticism of this method

was that it assumed the equivalence of sensory unpleasantness with ego unpleasantness.

3. Recall of experiences: many of these experiments have assumed actual numbers of pleasant and unpleasant experiences to be equal.

4. Discrete associations: association tests have in general shown more pleasant associations.

5. Reaction time: these experiments have shown shorter reaction time with pleasant material.

6. Memory for word lists: this method, the most frequently used, has yielded nothing significant.

7. Frame of reference: these studies have shown a selective effect of attitude on memory.

8. Specific studies: these studies, specifically reviewed in the article, while showing a clearer conception of the problem, have failed to demonstrate repression.

All reviews and criticisms of the problem have agreed that no adequate test of repression has been made. In most studies the repression has been conceived as differential forgetting. The most significant criticism has not been specifically formulated, namely, that *no test of repression can be considered adequate until the removal of the repression factor has resulted in the restoration to consciousness of the repressed material*. Any experiment which does not include this crucial step is not complete and the results can be attributed to other factors such as set, differential learning, differential motivation, practice, etc. rather than to active repression.

A proposed experimental design to test the hypothesis that repression proper is a process which inhibits recall without, however, destroying the memory for the subject matter, and that the memory can be restored by the elimination of the repression factor is presented below.

Control		Experimental	
I. Step 1. Learning	=	Learning	
Step 2. Retention test	=	Retention test	
Time interval			
II. Step 3. Retention test	=	Retention test	
Step 4. Neutral task		Repression	
Step 5. Retention test	>	Retention test	
Time interval			
III. Step 6. Retention test	>	Retention test	
Step 7. Neutral task		Removal of repression	
Step 8. Retention test	=	Retention test	

If the hypothesis is supported, the results indicated in the design can be expected. Any other results, although not disproving repression, would cast real doubt on the validity of the concept.

The design involves two equated groups, a control and an experi-

mental group. The individuals in each group are exposed to some material (Step 1) which they are to learn. Their degree of mastery is then measured with a retention test (Step 2). This satisfies the first requirement of an experimental test of repression proper, namely that the material be learned.

After a time interval another test of retention of the originally learned material is administered to the individuals in both groups (Step 3). Up to this point the performance of the groups should show no significant difference. At this point the procedure varies. The individuals in the experimental group are subjected to an ego threat associated with some activity while the individuals in the control group perform the same activity with no induced affect (Step 4). After this another retention test of the originally learned material is administered (Step 5). If the average retention in Step 5 is significantly better for the control group, the second criterion of an adequate experimental test of repression has been met.

After another time interval the subjects are again asked for a measure of retention of the originally learned material. (Step 6). This is followed by the same activity associated with the ego threat in Step 4, only this time there is no ego threat for either the experimental or the control group. Restoration of memory following the change in ego value of the task would constitute the third and crucial step in the experimental demonstration of repression proper.

BIBLIOGRAPHY

1. ANDERSON, A. C., & BOLTON, F. J. The inhibition of the unpleasant. *J. abnorm. soc. Psychol.*, 1925, 20, 300-302.
2. BALKEN, E. R. Affective, volitional and galvanic factors in learning. *J. exp. Psychol.*, 1933, 16, 115-128.
3. BARTLETT, F. C. *Remembering: a study in experimental and social psychology*. New York: Macmillan; Cambridge, Eng.: Univ. Press, 1932.
4. BAXTER, M. F., YAMADA, K., & WASHBURN, M. F. Directed recall of pleasant and unpleasant experiences. *Amer. J. Psychol.*, 1919, 30, 300-302.
5. BIRNBAUM, K. Ueber den Einfluss von Gefühlsfaktoren auf die Assoziation. *Monatsch. f. Psychiat. u. Neurol.*, 1912, 32, 95-123.
6. BUNCH, M. E., & WIENTGE, E. The relative susceptibility of pleasant, unpleasant, and indifferent material to retroactive inhibition. *J. exp. Psychol.*, 1933, 9, 157-178.
7. CARTER, H. D. Effect of emotional factors upon recall. *J. Psychol.*, 1936, 1, 48-55.
8. CASON, H. The learning and retention of pleasant and unpleasant activities. *Arch. Psychol., N. Y.*, 1932, No. 134. Pp. 96.
9. CHANEY, RUTH, & LAUER, A. R. The influence of affective tone on learning and retention. *J. educ. Psychol.*, 1929, 20, 287-291.

10. COLGRAVE, F. W. Individual memories. *Amer. J. Psychol.*, 1898-99, 10, 228-255.
11. EDWARDS, A. L. Political frames of reference as a factor influencing recognition. *J. abnorm. soc. Psychol.*, 1941, 36, 34-61.
12. EDWARDS, A. L. The retention of affective experiences—a criticism and restatement of the problem. *Psychol. Rev.*, 1942, 49, 43-53.
13. FLANAGAN, D. E. The influence of emotional inhibition on learning and recall. Unpublished Master's thesis, Univ. Chicago, 1930.
14. FLUGEL, J. C. A quantitative study of feeling and emotion in everyday life. *Brit. J. Psychol.*, 1925, 15, 318-355.
15. FOX, C. The influence of subjective preference on memory. *Brit. J. Psychol.*, 1922-23, 13, 398-405.
16. FRANK, J. P., & LUDVIGH, E. J. The retroactive effect of pleasant and unpleasant odors on learning. *Amer. J. Psychol.*, 1931, 43, 102-108.
17. FREUD, S. Repression. In *Collected papers*, Vol. 4, 84-97. London: Hogarth Press, 1925.
18. GILBERT, G. M. The age difference in the hedonistic tendency in memory. *J. exp. Psychol.*, 1937, 21, 433-441.
19. GILBERT, G. M. The new status of experimental psychology on the relationship of feeling to memory. *Psychol. Bull.*, 1938, 35, 26-35.
20. GORDON, KATE. Ueber das Gedachtnis fur Affective bestimmte Eindrucke. *Arch. ges. Psychol.*, 1905, 4, 437-458.
21. GORDON, KATE. Recollection of pleasant and unpleasant odors. *J. exp. Psychol.*, 1925, 8, 225-239.
22. GORDON, KATE. A study of early memories. *J. Delinqu.*, 1928, 12, 127-132.
23. GOULD, R. Repression experimentally analyzed. *Character and Pers.*, 1942, 10, 259-288.
24. GRIFFITHS, C. H. Results of some experiments on affection, distributions of associations and recall. *J. exp. Psychol.*, 1920, 3, 447-464.
25. HENDERSON, E. N. Do we forget the disagreeable? *J. Phil. Psychol. sci. Meth.*, 1911, 8, 432-438.
26. HUSTON, P. E., SHAKOW, D., & ERICKSON, M. H. A study of hypnotically induced complexes by means of the Luria technique. *J. gen. Psychol.*, 1934, 11, 65-97.
27. JERSILD, A. Memory for the pleasant as compared with the unpleasant. *J. exp. Psychol.*, 1931, 14, 284-288.
28. KENNETH, J. H. An experimental study of affects and associations due to certain odors. *Psychol. Monog.*, 1927, No. 37. Pp. 64.
29. KOCH, H. L. The influence of some affective factors upon recall. *J. gen. Psychol.*, 1930, 4, 171-190.
30. KOWALEWSKI, A. *Studien zur Psychologie des Pessimismus*. Wiesbaden, 1904.
31. KOWALEWSKI, A. *Schopenhauer und sein Weltanschauung*. Berlin, 1908.
32. LAIRD, D. The influence of likes and dislikes on memory as related to personality. *J. exp. Psychol.*, 1923, 6, 294-300.
33. LANIER, L. H. Memory for words differing in affective value. *Psychol. Bull.*, 1940, 37, 492-493. (Abstract.)
34. LEVINE, J. M., & MURPHY, G. The learning and forgetting of controversial material. *J. abnorm. soc. Psychol.*, 1943, 38, 507-517.
35. LURIA, A. R. *The nature of human conflicts*. (Trans. by W. H. Gantt.) New York: Liveright, 1932.
36. LYNCH, C. A. The memory value of certain alleged emotionally toned words. *J. exp. Psychol.*, 1932, 15, 298-315.
37. MCGEOCH, J. O. *The psychology of human learning*. New York: Longmans, Green, 1946.
38. MCGRANAHAN, D. V. A critical and experimental study of repression. *J.*

- abnorm. soc. Psychol., 1940, 35, 212-225.
39. McMULLIN, T. E. A study of the affective nature of the interpolated activity as a factor in producing different relative amounts of retroactive inhibition in recall and recognition. *J. exp. Psychol.*, 1942, 30, 201-215.
40. MELTZER, H. The forgetting of pleasant and unpleasant experiences in relation to intelligence and achievement. *J. soc. Psychol.*, 1930, 2, 217-227.
41. MELTZER, H. Individual differences in forgetting pleasant and unpleasant experience. *J. educ. Psychol.*, 1930, 21, 399-409.
42. MELTZER, H. The present status of experimental studies of the relation of feeling to memory. *Psychol. Rev.*, 1930, 37, 124-139.
43. MELTZER, H. Sex differences in forgetting pleasant and unpleasant experiences. *J. abnorm. soc. Psychol.*, 1931, 25, 450-464.
44. MENZIES, R. The comparative memory value of pleasant, unpleasant and indifferent experiences. *J. exp. Psychol.*, 1936, 18, 267-279.
45. MOORE, E. H. A note on the recall of the pleasant vs. the unpleasant. *Psychol. Rev.*, 1935, 42, 214-215.
46. MORGAN, E., MULL, H. K., & WASHBURN, M. F. An attempt to test moods or temperaments of cheerfulness and depression by directed recall of emotionally toned experiences. *Amer. J. Psychol.*, 1919, 30, 302-304.
47. MYERS, G. C. Affective factors in recall. *J. Phil. Psychol. sci. Meth.*, 1915, 12, 85-92.
48. O'KELLY, L. I., & STECKLE, L. C. The forgetting of pleasant and unpleasant experiences. *Amer. J. Psychol.*, 1940, 53, 432-434.
49. PETERS, W. Gefühl und Erinnerung; Beiträge zur Erinnerungsanalyse. *Psychol. Arbeit.*, 1911, 6, 197-260.
50. PETERS, W., & NEMECEK, O. Massenversuche über Erinnerungassoziationen. *Fortschr. Psychol. Anwend.*, 1914, 2, 226-245.
51. PINTNER, R., & FORLANO, G. The influence of pleasantly and unpleasantly toned words on retention. *J. soc. Psychol.*, 1940, 11, 147-148.
52. POSTMAN, L., & MURPHY, G. The factor of attitude in associative memory. *J. exp. Psychol.*, 1943, 33, 228-238.
53. PRENTICE, W. C. H. The interruption of tasks. *Psychol. Rev.*, 1944, 51, 329-340.
54. RAPAPORT, D. *Emotions and memory*. Baltimore: Williams and Wilkins, 1942.
55. RATLIFF, M. M. The varying function of affectively toned olfactory, visual and auditory cues in recall. *Amer. J. Psychol.*, 1938, 51, 695-699.
56. ROSENZWEIG, S. The recall of finished and unfinished tasks as affected by the purpose with which they were performed. *Psychol. Bull.*, 1933, 30, 698. (Abstract.)
57. ROSENZWEIG, S. The experimental study of repression. In H. Murray (Ed.), *Explorations in personality*. New York: Oxford Univ. Press, 1938. Pp. 472-491.
58. ROSENZWEIG, S. The experimental measurement of types of reaction to frustration. In H. Murray (Ed.), *Explorations in personality*. New York: Oxford Univ. Press, 1938. Pp. 585-599.
59. ROSENZWEIG, S. Need-persistent and ego-defensive reactions to frustration as demonstrated by an experiment on repression. *Psychol. Rev.*, 1941, 48, 347-349.
60. ROSENZWEIG, S. An experimental study of "repression" with special reference to need-persistent and ego-defensive reactions to frustration. *J. exp. Psychol.*, 1943, 32, 64-74.

61. ROSENZWEIG, S., & MASON, G. An experimental study of memory in relation to the theory of repression. *Lit. J. Psychol.*, 1934, 24, 247-265.
62. ROSENZWEIG, S., & SARASON, S. An experimental study of the triadic hypothesis: reaction to frustration, ego-defense, and hypnotizability. I. Correlational approach. *Character and Pers.*, 1942, 11, 1-19.
63. SARASON, S., & ROSENZWEIG, S. An experimental study of the triadic hypothesis: reaction to frustration, ego-defense, and hypnotizability. II. Thematic Apperception Approach. *Character and Pers.*, 1942, 11, 150-165.
64. SEARS, R. R. Functional abnormalities of memory with special reference to amnesia. *Psychol. Bull.*, 1936, 33, 229-274.
65. SEARS, R. R. An experimental test of one phase of the hypothesized repression sequence. *Psychol. Bull.*, 1936, 33, 744. (Abstract.)
66. SEARS, R. R. Initiation of the repression sequence by experienced failure. *J. exp. Psychol.*, 1937, 20, 570-580.
67. SEARS, R. R. Non-aggressive reactions to frustration. *Psychol. Rev.*, 1941, 48, 343-348.
68. SEARS, R. R. Survey of objective studies of psychoanalytic concepts. *Soc. Sci. Res. Coun. Bull.* No. 51, 1943.
69. SEARS, R. R. Experimental analysis of psychoanalytic phenomena. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*. New York: Ronald Press, 1944. Pp. 306-332.
70. SEELEMAN, V. The influence of attitude upon the remembering of pictorial material. *Arch. Psychol.*, N. Y., 1940, 36, No. 258.
71. SHARP, AGNES A. The influence of certain emotional inhibitions on learning and recall. Unpublished Master's thesis, Univ. Chicago, 1930.
72. SHARP, AGNES A. An experimental test of Freud's doctrine of the relation of hedonic tone to memory revival. *J. exp. Psychol.*, 1938, 22, 395-418.
73. SHAW, F. J. Two determinants of selective forgetting. *J. abnorm. soc. Psychol.*, 1944, 39, 434-445.
74. SHAW, F. J., & SPOONER, A. Selective forgetting when the subject is not ego involved. *J. exp. Psychol.*, 1945, 35, 242-247.
75. SILVERMAN, A., & CASON, H. Incidental memory for pleasant, unpleasant and indifferent words. *Amer. J. Psychol.*, 1934, 46, 315-320.
76. SMITH, W. W. Experiments on memory and affective tone. *Brit. J. Psychol.*, 1921, 11, 236-250.
77. STAGNER, R. The reintegration of pleasant and unpleasant experiences. *Amer. J. Psychol.*, 1931, 43, 463-468.
78. STAGNER, R. Factors influencing the memory value of words in a series. *J. exp. Psychol.*, 1933, 16, 129-137.
79. STECKLE, L. C. Again—affect and recall. *J. soc. Psychol.*, 1945, 22, 103-105.
80. STONE, A. R. The reaction of memory to affective states. *Amer. J. Psychol.*, 1925, 36, 112-123.
81. SUSUKITA, T. Ueber das Gedächtnis für lust- und unlustbetonte Erlebnisse im Alltagsleben. *Tohoku Psychol. Folia*, 1935, 3, 187-204.
82. TAIT, W. D. Effects of psychophysical attitudes on memory. *J. abnorm. soc. Psychol.*, 1913-14, 8, 10-38.
83. THOMPSON, R. H. An experimental study of memory as influenced by feeling tone. *J. exp. Psychol.*, 1930, 13, 462-467.
84. TOLMAN, E. C. Retroactive inhibition as affected by conditions of learning. *Psychol. Monogr.*, 1918, No. 107, 187-195.
85. TOLMAN, E. C., & JOHNSON, I. A note on association time and feeling.

- Amer. J. Psychol.*, 1918, 29, 187-195.
86. WALLEN, D. Ego involvement as a determinant of selective forgetting. *J. abnorm. soc. Psychol.*, 1942, 37, 20-29.
87. WATERS, R. H., & LEEFER, R. The relation of affective tone to the retention of experiences in everyday life. *J. exp. Psychol.*, 1936, 19, 203-215.
88. WATSON, W. S., & HARTMANN, G. W. The rigidity of a basic attitudinal frame. *J. abnorm. soc. Psychol.*, 1939, 34, 314-335.
89. WHITE, M. M. Some factors influencing recall of pleasant and unpleasant words. *Amer. J. Psychol.*, 1936, 48, 134-139.
90. WHITE, M. M., & RATLIFF, M. M. The relation of affective tone to learning and recalling words. *Amer. J. Psychol.*, 1934, 46, 92-98.
91. WOHLGEMUTH, A. The influence of feeling on memory. *Brit. J. Psychol.*, 1923, 13, 405-416.
92. ZEIGARNIK, B. Über das Behalten von erledigten und unerledigten Handlungen. *Psychol. Forsch.*, 1927, 9, 1-85.
93. ZILLIG, M. Einstellung und Aussage. *Z. Psychol.*, 1928, 106, 58-106.

Received June 30, 1949.

ON THE CIRCULARITY OF THE LAW OF EFFECT

PAUL E. MEEHL

University of Minnesota

In his recent review on "The History and Present Status of the Law of Effect," Postman (19) lays considerable emphasis on the problem of "circularity" which he sees as crucial in the formulation of the law. He says:

Whereas some critics were most concerned with the mechanisms mediating effect, others focussed their attention on the nature of the satisfiers and annoyers to which reference is made in Thorndike's law. Although Spencer and Bain, in whose tradition Thorndike continued, frankly invoked pleasure and pain as agents responsible for the fixation and elimination of responses, Thorndike's law has been a law of *effect*, not *affect*. He carefully defines satisfiers and annoyers in terms independent of subjective experience and report. "By a satisfying state of affairs is meant one which the animal does nothing to avoid, often doing such things as to attain and preserve it. By a discomforting state of affairs is meant one which the animal avoids and abandons." Although admittedly free of hedonism, such a definition of satisfiers and annoyers has faced another serious difficulty: the danger of circularity. The critic may easily reword the definition to read: "The animal does what it does because it does it, and it does not do what it does not do because it does not do it." This *reductio ad absurdum* is probably not entirely fair, but it points up the danger of the definition in the absence of an *independent* determination of the nature of satisfiers and annoyers. The satisfying or annoying nature of a state of affairs can usually be determined fully only in the course of a learning experiment and cannot then be invoked as a causal condition of learning without circularity. In their experimental work Thorndike and his associates have made no significant attempts to establish the satisfying or annoying nature of their rewards and punishments independently of the learning experiment (19, p. 496).

And a little later Postman says:

Stripped of virtually all defining properties and qualifications, the law does indeed have a very wide range of applicability but only at the expense of vagueness. The sum and substance of the argument now is that something happens in the organism (nervous system) after an act is performed. The fact that something happens influences further action. This something is, however, so little defined that it has almost no predictive efficiency. The O.K. reaction has no measurable properties, the conditions for its occurrence are so general as to embrace almost every conceivable situation. Hence the operation of O.K. reaction can be inferred only *ex post facto*, after learning has taken place. But here we are again impaled on the horns of the dilemma of circularity (19, p. 497).

And still further:

In attempting to evaluate the controversy which has raged around the

definition of satisfiers one is struck by the key importance of the hedonistic issue. Certainly hedonism is an immediate ancestor of the law, and now that the principle of effect has reached an uneasy maturity it is clear that it cannot deny its origin without sacrificing much of its vigor. When the law is stripped of hedonistic implications, when effect is not identified with tension-reduction or pleasure (as by Thorndike), the law of effect can do no more than claim that the state of affairs resulting from a response in some way influences future responses. Such a statement is a truism and hardly lends itself to the rigorous deduction of hypotheses and experimental tests. If a neohedonistic position is frankly assumed (as, e.g., by Mowrer) the law becomes an important tool for research, provided "satisfaction" is independently defined and not merely inferred from the fact that learning has occurred (19, p. 501).

Throughout Postman's paper this problem is constantly lurking behind the scenes even when the author does not single it out for specific mention. I am in complete agreement with Postman's final remark that "at the present state of our knowledge the law of effect as a monistic principle explaining all learning has not been substantiated," and Postman performs a service by emphasizing this problem of circularity in his discussion of the "law." I am inclined, however, to think that he has settled the question of circularity somewhat too easily, and that his settlement of it has an effect upon much of his argumentation. I gather from the above quotations that Postman looks upon any definition of effect or reinforcement in terms of the resulting change in response strength as "circular," where that word has a pejorative sense. If he is right in this it is very serious. While the law of effect has many difficulties, I do not believe that "circularity" is among them. To show this is the aim of the present paper.

I shall consider the problem of circularity in the law of effect as identical with the problem of circularity in the definition of *reinforcement* in instrumental conditioning. I take it that Postman does the same, since in the first quotation above he cites a passage from Hilgard and Marquis' *Conditioning and Learning*, where the two problems are considered together and with free interchange of the two terminologies. These authors say:

It is apparent that no definition of effect provides an independent measure of the strength of reinforcement. The degree of satisfaction, of complacency, or of tension reduction has not been objectively determined. The strength of reinforcement can be given comprehensive definition only in terms of the amount of learning resulting from it. This is, of course, a circular definition, if strength of reinforcement is to be used as a factor determining degree of learning. A partial escape from circularity is achieved by the fact that a stimulus such as food which is found to be reinforcing in one situation will also be reinforcing in other situations, and with other animals (9, p. 83).

Writing in 1948, however, Hilgard states concerning Thorndike's "operational" definition of satisfiers and annoyers:

These definitions are not circular, so far as the law of effect is concerned. That is, the states of affairs characterized as satisfying and annoying are specified independently of their influence upon modifiable connections. The law of effect then states what may be expected to happen to preceding modifiable connections which are followed by such specified states. The objection that Thorndike was lacking in objectivity in the statement of the law of effect is not a valid one (8, p. 24).

Hilgard is willing to let the concept of reinforcement (effect, satisfaction, reward) be introduced on the basis of behavior, but only because there are behavioral criteria of seeking and avoiding other than the effect of reinforcement upon *modifiable* connections. Whether this restriction is necessary needs to be considered carefully.

Skinner dismisses the whole problem in two sentences:

A reinforcing stimulus is defined as such by its power to produce the resulting change. There is no circularity about this; some stimuli are found to produce the change, others not, and they are classified as reinforcing and nonreinforcing accordingly (22, p. 62).

Spence (23) takes essentially the same tack in his recent discussions of secondary reinforcement. The stimuli which impinge upon an organism may be divided, he says, into two classes: those which produce an increment in response strength, and those which do not. It seems from the several preceding quotations that there is a lack of agreement as to whether or not the law of effect or the principle of reinforcement involves an unavoidable circularity, or, if it does not, how circularity is to be avoided. In what follows, I make no claim to originality, since the essence of my development is contained in the previous quotations, together with the work of Tolman. But I feel it worthwhile to bring the arguments together in one context, and to show that the problem merits somewhat more extended treatment than is usually given it. Without claiming to present a definitive solution, I shall indicate the general direction which I believe the solution might take, and in the process introduce certain distinctions and terminological proposals which I feel might clarify our discussion and experimentation.

THE MEANING OF CIRCULARITY

It must be pointed out that there are two meanings of the word "circular" in common use. We have on the one hand circularity in *definition*, in which an unfamiliar term is defined by using other terms which are (directly or ultimately) defined by the term in question. There is no question of circularity in this sense in a definition of the

Skinner-Spence type. Let us accept as a crude preliminary formulation the following: "A reinforcing stimulus is one which increases the subsequent strength of responses which immediately precede it." The words *stimulus*,¹ *strength*, *increase* and *response* are all definable without any reference to the fact or theory of reinforcement. The definitions of these terms, particularly the term "response," present terrible difficulties; but I do not know of anyone who maintains that they involve the notion of reinforcement. Words such as these are current in the vocabulary of many kinds of psychological theorists who do not accept the Law of Effect as a principle of learning and in the absence of any indications to the contrary, I shall assume that we can tell what we mean by them. We can determine empirically when the strength of a response has increased without knowing anything about reinforcing stimuli, drives, satisfactions, and the like. It seems clear that the definition of a reinforcing stimulus in terms of its effect on response strength does not involve circularity in *this* sense.

The other meaning of the word circularity refers not to meanings (definition of terms) but to the establishment of propositions. We speak of *proofs* as being circular if it can be shown that in the process of establishing (proving) a proposition we have made use of the probandum. I am not aware that any responsible theorist has attempted to "prove" the Law of Effect in this way. It is true that those who look upon the law as fundamental are skeptical when they hear of a case of increase of response strength which does not *seem* to involve any obvious reinforcing consequences so that they begin to invent hypotheses to explain the results. There is no harm in this so long as the proposed explanations are in principle confirmable on the basis of some other experimental consequences, however remote. If an animal learns a response sequence without being given food, water, or any of the usual rewards, I suspect most Hullians would begin to talk about secondary reinforcement present in the situation. One can, of course, be careless with this kind of explanation, but there is nothing intrinsic to the concept that entails non-confirmability. The establishment of secondary reinforcing effects as explanations of a given experimental result consists in combining the facts known about primary reinforcers with facts about the animal's life history, in terms of which we understand how certain stimuli have acquired their secondary reinforcing powers. People on both sides of the present controversy over reinforcement theory are performing many different sorts of experiments in

¹ "Stimulus" will be used broadly to include "stimulus change," and stimulus configurations of all degrees of patterning and complexity.

order to confirm or disconfirm the Law of Effect. It would seem that if the law of effect *were* being treated by anyone as a consequence of definition, or established by some hidden assumption of its truth, the experiments would not be going on.

CAN "REINFORCEMENT" BE INDEPENDENTLY DEFINED?

Nonetheless, when we think about this definition we feel uncomfortable. I do not think we have in mind either a circularity in definition or a begging-the-question fallacy, but some sort of peculiar pseudo-circularity in which it seems to us vaguely that the law *could* be "derived" from the proposed definition, even though no one in fact seems to be trying to do it this way. The problem can be stated very simply: How can we introduce the concept of reinforcement in terms of effect upon strength, and still have a "law of effect" or "principle of reinforcement" which has the empirical content that everybody seems to be taking for granted in experimentation?

1. Suppose we reject the Thorndike-Skinner-Spence procedure of defining reinforcement in terms of response strength, and decide to define the term quite independently of the learning process. The first possibility, which we shall dismiss rather dogmatically, is to do it subjectivistically in terms of pleasure, experiences of satisfaction, and the like. Aside from the general behavioristic objections, and the specific problems of measurement created, this approach is not feasible because it leaves us without any basis for speaking of reinforcing value in the case of that very important class of motivations that are unconscious or at least inadequately verbalized in the human case; and it makes impossible the establishment of reinforcing value in the case of lower organisms. At the present time there are probably very few psychologists who would consider this alternative seriously.

2. Secondly, we might try to define reinforcers in terms of certain physical properties on the stimulus side. I shall attempt to show below that this is a procedure which *follows* the introduction of the generic notion of a reinforcer, and which at a later stage becomes very important. But no one wants to group together an arbitrary class of physical objects or stimuli and call them "reinforcers," since the aim of our concept formation is to make possible the statement of laws. The possibility of identifying common physical properties of that large class of stimuli already grouped together as "rewarding" seems very remote. Besides, we would set up these properties or sets of properties by examining the members of the reinforcing class, which we already have set apart on some basis or other; and the question is: How have we arrived at the members of that class?

3. A third possibility, seen in the work of Hull, is to define reinforcement ultimately in terms of drive reduction, that is, in terms of the inner physiological events involved. Here again, I do not suppose that anyone would be able to give even the vaguest specification of the defining property of all neural events which are reinforcing. Even for the so-called primary physiological needs such as hunger, the evidence as to their exact physiological basis is most

incomplete. No psychologist today is willing to equate "hunger" with "stomach contractions," in the light of the experimentation on visceral denervations, specific sub-hungers, and the like. In other cases, we have practically no information on the neurophysiology, e.g., the neurophysiologic basis of the reinforcing effect of the presence of another organism, the turning off of a light in the Skinner box, or the going through of "exploratory" behavior on the other side of a grill. There is some reason to suppose that certain stimuli retain their secondary reinforcing value in the absence of the primary drive (2, 16), which complicates the problem further.

These considerations force a return to the *effect* of stimuli as a basis for specifying that they are reinforcers, and this leads to the paradox. If we define a reinforcing agent by its effect upon learning, then it seems that whenever learning is effected, we know ("by definition") that we have given a reinforcement. For surely, when the organism behaves, some stimulus change occurs, if nothing else than the proprioceptive effects of responding. If the behavior increases in strength, then these stimulus changes, which were in fact preceded by the response, are reinforcers. Hence, it seems that a definition of reinforcement in terms of an increase of habit strength makes the law tautological and devoid of factual content. This train of thought, which I am sure is familiar to most readers, seems obvious and straightforward. But I believe it can be shown to be mistaken, once the law is stated *explicitly* in the way we all really think of it *implicitly* when we perform experiments or try to explain a given case of learning.

AN EMPIRICAL DERIVATION OF REINFORCEMENT

Let us begin afresh by going to the behavior itself in a situation in which there is little or no disagreement as to what occurs. Consider a bright, inductively inclined Martian, who had never experienced any needs or satisfactions (except perhaps a *Cognizance!*) and who was observing the behavior of a rat in successive runnings in a T-maze. For the moment we shall simply consider a "standard rat," neglecting the individual differences in parameters and the accidents of personal histories that generate special secondary reinforcing properties. These refinements need to be added later, but as is usually the case will have to be added by being integrated into the whole structure of reinforcement theory, since we cannot treat everything at once. At the beginning, the Martian observes that the organism turns to the right or left with, let us say, about equal frequency. With further trials, a change occurs until finally the rat is responding close to 100% of the time by turning to the right. A Martian could obviously discover this with no notion of rewards, pleasure and the like. If he is ingenious enough to think of the possibility that the strength of a response might be influenced by the events that follow it in time, he would then proceed to

investigate the changes that are contingent on this right turning.³ He notes that when the rat turns to the right he brings about the following states of affairs on the stimulus side which he does not bring about when he turns to the left: He ends up nearer to the right-hand wall, which is painted green; he twists his own body to the right in responding; he ends up in a wooden box having knots in the wood; he ends up nearer the North pole; and to a dynamo on the other side of the campus; and he comes into the presence of a cup of sunflower seeds. These are the stimuli (stimulus changes) which are contingent on right turns. Is it possible that the gradual strengthening of the right turning is dependent upon one, some, or all of these changes following it? Our scientist from Mars would proceed to study a series of standard rats in the situation, altering the above variables systematically by usual inductive procedures. As a matter of empirical fact, he would discover that, within certain very wide limits, alterations in the first five have no effect. The sixth, the sunflower seeds, have a tremendous effect. He finds that he can alter the geographical direction, the direction of the body twist required, the wall color approached, etc.—that he can introduce all manner of modifications in the other factors; and so long as the sunflower seeds are presented, the rat will tend to go to where they are. On the other hand, if the sunflower seeds are omitted, and nothing else put in their place, a preference fails to develop as a function of these remaining differences.

But we have already greatly over-simplified. Actually, the Martian would discover that the effect of finding sunflower seeds in some cases is almost too slight to be detected; furthermore, even after a preference has been acquired, it may on some occasions fail to show itself. Now, it has already been apparent that when he comes upon these sunflower seeds, the rat behaves toward them in a characteristic way, that is, he ingests them. In seeking to understand the variability in the development and manifestation of a preference, one would notice a correlation between the strengthening of a preference and the rate, strength, and consistency of ingestive responses in the presence of the food. Identifying the same rat on successive days, it is found that on those days on which a preference already established broke down, very frequently the ingestive response in the presence of the sunflower seeds was at a very low or even zero strength. Failing to find anything varying in the maze situation itself to account for these differences, one can study the expe-

³ Actually, no great ingenuity is involved here. Study of the events immediately *preceding* a run, e.g., the manner in which the experimenter handles the rat, what orientation he gives its head in placing it in the entry box, etc., would fail to reveal any systematic factor related to the direction of a preference. Considering this, together with the fact that before any runs have been made no preference exists, the Martian would be led to ask whether it is something that happens *after* the run (or during it) that affects the probability of a similar choice in subsequent runs.

riences of the animals between runs. Here appears a very striking correlate of both preference strength *and* the ingestive response in the maze: that which a human experimenter would call the "feeding schedule." The Martian would observe that when sunflower seeds were made available to the rats in their cages, they behave with respect to them in the same way as they do when they come upon the sunflower seeds in the goal box: namely, with ingestive responses. He would discover, again by systematic variation in these conditions, that such matters as the chemical nature of the substance made available, the periodicity of its availability, the lapse of time between when it was last available and the maze run; the rate of ingestion manifested at the moment of beginning deprivation (i.e., how close the rat was to satiety when interrupted), and so on, all exert an effect upon the maze response. By far the most intimate correlate would be the lapse of time since feeding. To quote Skinner again,

The problem of drive arises because much of the behavior of an organism shows an apparent variability. A rat does not always respond to food placed before it, and a factor called its "hunger" is invoked by way of explanation. The rat is said to eat only when it is hungry. It is because eating is not inevitable that we are led to hypothesize the internal state to which we may assign the variability. Where there is no variability, no state is needed. . . . In dealing with the kind of behavior that gives rise to the concept of hunger, we are concerned with the strength of a certain class of reflexes and with the two principal operations that affect it—feeding and fasting (22, pp. 341, 343).

For a considerable class of stimuli found to affect choice behavior in the maze, there is a fairly well demarcated class of events in the extra-maze activities which exert an effect. Food, water, a female rat, all depend for their efficacy upon a deprivation schedule of some sort. For other stimuli, the rest of the day's activities seem of less relevance. For example, the effects of turning off a light in the Skinner box upon the lever pressing response would not depend upon a schedule of extra box illumination in any such obvious way as the effects of a food pellet depend upon the extra maze operations of feeding and fasting. Even here, at the extremes, it is likely that the schedule has some effect. Although I know of no experimental material on the point, it would be surprising if rats raised and maintained in a dark or extremely bright living cage would show the same response to light-off as a reinforcing agent. In order to keep the discussion quite general, I shall refer to *schedule-reinforcer* combinations, which will be understood to include those combinations in which almost any life-maintaining schedule is adequate. Whether there are any such does not need to be settled here. The stimulus presented is a *reinforcer*, and the presentation of it (an "event") is a *reinforcement*.

We are now in possession of a rather simple set of empirical facts. A certain stimulus, for a rat which has been under a specified schedule,

for instance the sunflower seeds for a rat who has not ingested anything for 23 hours, will exert a strengthening effect. We can formulate a "law" stated crudely as follows: "In a rat which has not recently ingested sunflower seeds, bran mash, Purina chow, etc., a response of turning in a given direction in the T-maze will be increased if the fairly immediate presentation of sunflower seeds, etc., is made contingent upon that response." Similarly, we would find such a specific law to hold for thirst and water, sex and a mate, and so on. The general form of such special laws would be: "On schedule M, the termination of response sequence R, in setting S, by stimulus S^1 is followed by an increment in the strength of S.R." Such a law may be called a *situational-reinforcement* law, where the "reinforcement" is understood to stand for "presentation-of-a-reinforcer-following-a-specified-maintenance-schedule," and the term "situational" covers "response R in situation S."

Actually, in any given case, M, R, S, S^1 are classes. This is indicated by the suspicious-looking "etc." in the first "law" above. There is nothing shady about this "etc.," inasmuch as what is actually involved here is a class of operations and effects which are ultimately to be specified by locating each instance with respect to a whole complex set of dimensions. For example, Guttman (6) shows a relation between concentration of sugar solution used as a reinforcing agent and the strength of the lever pressing response. Heron and Peake (7) have studied protein as a specific component of reinforcement. There is to be discovered a vast number of such rather special laws which are comparable to the myriads of laws in chemistry concerning the solubility of substance Y in substance X and the like.

The next thing to notice is that while the schedule, reinforcement, response, and situation are all classes showing certain relations to one another, in general the schedule and reinforcer are related to one another more intimately than either is to the situation or response. The strength of a response which is maintained by food reinforcement is heavily dependent upon the feeding-fasting schedule, whereas the effect of a food reinforcement upon a response is relatively independent of, say recency of copulatory activity, so that a given schedule-reinforcement pair are "tied" to one another. But the Martian observes that the strengthening effect of a given schedule-reinforcement combination is relatively (not wholly!) neutral with respect to the response we are trying to strengthen and the situation in which we are trying to strengthen it. For a hungry rat, right turning depends heavily upon finding food; for a satiated rat, it depends very little. So the feeding schedule is intimately related to the reinforcing agent's efficacy. However, this "hungry-food" schedule-reinforcement combination seems to be capable of strengthening chain-pulling, lever-pressing, wheel-turning, marble-rolling, gnawing-through-paper, and so on through a very wide range of behaviors differing greatly in their topography and in their

stimulus conditions. This leads to the question, will a certain schedule-reinforcer combination increase the strength of *any* response, in *any* setting?" This question turns out empirically to be answered in the negative, since we find at least three limitations upon the generality of a schedule-reinforcer combination as response strengthener. Leaving out the trivial case in which the response is anatomically impossible, e.g., to teach an elephant to thread a needle, we find:

1. No situation-response sequences may involve stimulus dimensions which are not discriminable by the organism. (Tolman's "discriminating capacities").
2. Some response sequences seem on the basis of their sequence, timing, or "complexity" not to be learnable by members of a given species, or subgroups within a species. It appears impossible to teach a rat a quintuple alternation problem, or to teach a human moron integral calculus.
3. There are cases in which the response we wish to strengthen is incompatible with responses at a very high (and relatively unmodifiable) strength under the schedule-stimulus combinations we are employing. For example, it would probably be next to impossible to teach a very hungry cat to carry a piece of fresh liver across the room, deposit it in a box, and return to receive food as a reinforcement. "Defensive" and "anxiety-related" responses are among the most important examples of this case.

How do we discover what responses have these characteristics? Experimentally, as we discover anything else. Let us call a situation-response combination having none of these properties *learnable*. A positive definition will be given below. What we find is that whereas learnable responses seem to differ somewhat in their "readiness" under different schedule-reinforcement combinations, this is a matter of parameters and does not invalidate the following tentative "law," which is stated qualitatively: "Any learnable response will be strengthened by sunflower seeds as a reinforcer." The general form of such a law is "the stimulus S^1 on schedule M will increase the strength of any learnable response." I shall call such a law a *trans-situational reinforcement* law. It must be noted carefully that such a law is still about a *particular* reinforcing agent, having, to be sure, a class character; but the particular reinforcing agent (and its associated necessary schedule, if any) is no longer tied to the response sequence first studied. The reinforcing property of sunflower seeds was noted first in the T-maze. The Martian will discover that white rats *can* learn to pull chains, press levers, and roll marbles. He finds that these learnable responses can also be strengthened by making the feeding of sunflower seeds contingent upon them. He makes the inductive generalization that sunflower seeds would exert this effect upon all learnable responses in the rat.

He now asks the obvious question: Are all schedule-reinforcer combinations like this? That is to say, when we study a new schedule-reinforcer combination and find it strengthens a response, can we assume that it will increase the strength of all learnable responses?

Naturally, our confidence in the general reinforcing power of any particular one will increase as we try it out on more and more learnable responses. But we do not know whether a higher-order inductive statement is justified, so long as we study sunflower seeds only or study several kinds of agents but in only one situation each.

Having found a particular reinforcer in a particular situation, we have discovered that it is trans-situational. Next we discover that all of the reinforcers that we have investigated have turned out to be trans-situational. The next induction is, "If a learnable response is followed by a stimulus which is known to be a reinforcer of learnable responses the strength will increase." A shorter way of saying this, having first defined a reinforcer as "a stimulus which will increase the strength of at least one learnable response," is simply: *all reinforcers are trans-situational*. Nothing is said as to the *amount* of strengthening. It is sufficient, in order to demonstrate the trans-situational character of a reinforcing agent, to show that it produces an increment in strength. If equal increments were required, it is probable that very few (if any) reinforcers would be trans-situational because of the varying behavior readinesses and different parameters of habit acquisitions from one drive to another and from one situation to another.

This assertion, that all reinforcers are trans-situational, I propose to call the *Weak Law of Effect*. It is not our problem in this paper to discuss whether the Weak Law of Effect holds or not. A "proof" of the Weak Law of Effect consists, as usual, of establishing inductively many instances of it in a variety of situations with our confidence increasing on the basis of the usual inductive canons. A "disproof" of the Weak Law of Effect would involve showing that a certain stimulus change acts as a reinforcing agent for one response, i.e., that the presentation of this stimulus following the response will increase the latter's strength; but that another response, previously established as learnable, cannot be strengthened by a presentation of this agent. A failure of the Weak Law of Effect to hold strictly would not be particularly serious, since one could (at the very least!) specify the exceptions and would hope to be able to generalize about them, that is, to discover empirically what are the kinds of reinforcers, or kinds of differences among situations, which reveal its invalidity. Actually, here again we have a case in which the law is stated in a qualitative all-or-none form; but the development of a science of behavior would eventually result in substituting a multiplicity of laws indicating the extent to which the reinforcing (strengthening) property generalized over various dimensions of the stimulus side, the reinforcing agent, and the "required" response properties. Assuming the Weak Law of Effect to have been established inductively, where are we now in our development? We have specific situation-reinforcer laws which state that a given stimulus is a reinforcing agent for a specified kind of response in a specified situation. As an

example, we discover that for a standard rat, sunflower seeds will strengthen right turning in the T-maze. Having established several such specific situation-reinforcer laws, we find it convenient to introduce a definition, saying that a situational reinforcer is a stimulus which occurs as a term in such a specific situation-reinforcer law. Sunflower seeds are hence situational reinforcers. This definition is "arbitrary" or "conventional" in the usual sense, but clearly leads to no circularity. We cannot tell from the definition whether or not there is such a thing as a situational reinforcer, just as we cannot tell from the definition of a unicorn or of the phrase "King of France" whether such a thing exists. All we stipulate in the definition is that if a thing having certain properties turns out to exist, we will call it by this name. That there are situational reinforcers, that is to say, that we can find stimuli that do increase the strength of responses in a certain situation, is an empirical result. It is obvious that the specific situation-reinforcer laws have a perfectly good factual content (e.g., each such law could be false) in spite of the conventional character of the definition.

If our science contained nothing but a collection of such situational-reinforcer laws, we would still be in possession of valuable information. But we discover inductively that we can actually say more than this. For any given reinforcer, we discover that it can in fact be used to increase the strength of responses differing very greatly in topography from the one which originally led us to infer that it was a reinforcer, and in very different stimulating fields. It is true that there are a few special cases, as our cat with the liver, in which we cannot increase the strength of a *kind* of a response (carrying an object from one place to another) which we know from independent study this species is able to learn. But in all such cases we are able to specify an interfering response at such high strength that the behavior in question does not get a chance to be emitted, and hence cannot be reinforced. With this exception, we are able to say that a given reinforcer will increase the strength of all learnable responses of the species; although there will be quantitative differences which remain to be discovered and generalized about after much painstaking experimentation. We define a reinforcer which is of this sort as trans-situational, and from a study of numerous reinforcers we conclude that they are all of this type. The second order induction that all reinforcers are trans-situational (the Weak Law of Effect) is then made.

This last is certainly a very rich and powerful induction. It is true that to make predictions we must study at least one learnable response in order to find out whether a given stimulus change is reinforcing, and we must know for any contemplated response whether it is learnable. Experience with a given species need not be too extensive in order to get a general idea of the kinds of behavior which are possible and learnable; and once having this, we proceed to strengthen responses by means of

reinforcing agents which have never been utilized before in connection with these responses. This is so commonplace that we are likely to underestimate its theoretical significance. So far as I know, no animal psychologist has the least hesitation in utilizing any of a very large class of reinforcing objects called "food" in experimentation upon practically any kind of behavior which he is interested in studying. Should he find a failure of response strength to increase, the chances of his asking what is wrong with the food are negligible. His inductive confidence in the Weak Law of Effect is such that he will immediately begin to investigate what is wrong with the stimulus field, or what requirements concerning the response properties he has imposed which transcend the powers of the organism. I am stressing this point because there is a tendency to say that since we have to study the effects upon strength in order to know whether an agent is reinforcing, we do not really "know anything" when we have enunciated the Law of Effect. I think it should be obvious from the diversity of both the class called learnable and the class of agents called reinforcing that to the extent that this law holds almost without exception, when we have enunciated it we have said a great deal.

The man from Mars might be tempted here to take a final step which would be suggested by the ubiquity of the manifestations of the Weak Law of Effect. It might occur to him that the great majority of the instances in which changes in response strength occur seem to involve the operation of the Weak Law, i.e., the presentation of a member of the reinforcing class. Perhaps it is not only true that any learnable response can be strengthened by the presentation of a trans-situational reinforcer but may it not be that this is the *only* way to increase the strength of responses (by learning)? Response strength may be increased by surgical and drug procedures, and also by maturation; but the demarcation of learning as a very general mode of response change, while it presents difficult problems, need not concern us here. Assuming that we have some satisfactory basis for distinguishing an increase in the strength which is based upon "experience" rather than upon interference with the reaction mechanism or biological growth determined by genetic factors given minimal (viable) environments, we may ask whether learning takes place on any *other* basis than the Weak Law of Effect. Certain apparent exceptions to this statement of reinforcement as a necessary condition would appear, but the Martian might ask whether these exceptions are more apparent than real. The formulation of such a law would run something like this: "Every learned increment in response strength requires the operation of a trans-situational reinforcer." I shall designate this rash inductive leap as the *Strong Law of Effect*.

It appears obvious that this also is a statement far from being experimentally empty or in any sense a consequence of definition. I

have heard psychologists translate the statement "he learns because he was reinforced" as being tantamount to "he learns because he learns." Postman suggests the same kind of thing in the first quotation above. This is too easy. The expanded form which I suspect everyone has implicitly in mind when he talks about the Strong Law of Effect is: "He learns following the presentation of a stimulus change which for this species has the property of increasing response strength; and, other things being equal in the present setting, if this change had *not* occurred he would not have learned." Such a statement can clearly be false to fact, either because no such trans-situational reinforcer can be shown to have been present, or because the same learning can be shown to be producible without it in the present setting. The claim of the reinforcement theorist to explanation is (at this stage of our knowledge) of exactly the same character as "he developed these symptoms because he was invaded by the Koch bacillus, and we know that the Koch bacillus has these effects." This is not a very *detailed* explanation, because the intermediate or micro-details of the causal sequence are not given; but it is certainly neither factually empty nor trivial.

In our initial quotation from Postman, we find him saying, "The satisfying or annoying nature of a state of affairs can usually be determined fully only in the course of a learning experiment and cannot then be invoked as a causal condition of learning without circularity." The trouble with this remark lies in the ambiguity of the phrase "*a* learning experiment." That we cannot know what is reinforcing without having done *some* experimentation is obvious, and is just as it should be in an empirical science. But once having found that a certain state of affairs *is* reinforcing for a given species, there is no reason why a given case of learning cannot be explained by invoking the occurrence of this state of affairs as a causal condition. The definition of force does not entail the truth of Hooke's law. It is only by an experiment that we find out that strain is proportional to stress. Once having found it out, we are all quite comfortable in utilizing Hooke's law to account for the particular cases we come across. I am confident that Postman would not be disturbed if in answer to the question, "Why does that door close all the time?" someone were to reply, "Because it has a spring attached to it on the other side." There is no more "circularity" in this kind of causal accounting than in any other kind. It is perfectly true that this kind of "lowest-order" explanation is not very intellectually satisfying in some cases, although even here there is a considerable variability depending upon our familiarity with the situation. For a detailed consideration of these problems by more qualified persons I refer the reader to papers by Hospers (10), Feigl (4, 5), and Pratt (20).

I think it is obvious that this is the way we think of the Law of Effect, whatever we may think as to its truth. When an apparent case of learning in the absence of reinforcement occurs, those who are interested in preserving the status of the Law of Effect (in my terminology, in preserving the status of the *Strong* Law of Effect) begin to search for changes following the response which can be shown to be of the reinforcing sort. They do not simply look for *any* stimulus change and insist ("by definition") that it is a reinforcement. The statement that a given case of apparently non-reinforcement learning is actually based upon secondary reinforcement is essentially a claim that some stimulus change can be shown to have followed the strengthened response, and that this stimulus change has (still earlier) been put in temporal contiguity with a stimulus change of which we know, from a *diversity* of situations, that it exerts a reinforcing effect.

Abandoning the charge of circularity, a critic might offer a "practical" criticism, saying, "What good does it do to know that a reinforcer strengthens, when the only way to tell when something is a reinforcer is to see if it strengthens?" The trouble here lies in the vagueness, since the *generality* is not indicated, and this failure to indicate generality neglects the usual advantages of induction. That a describable state of affairs *is* reinforcing can only be found out, to be sure, by experimenting on some organisms utilizing *some* learnable response. But it is not required (if the Weak Law of Effect is true) that we, so to speak, start afresh with each new organism of the species and each new response. As a matter of fact, after we have considerable experience with a given species, we can generalize about the physical properties of a stimulus class. So that finally "food" means many substances which may never yet have been tried in a learning situation, and may never have been presented in natural circumstances to the members of a particular species. Wild rats do not eat Purina Chow. Here we begin to approach inductively one of the previously rejected bases of defining reinforcement, namely, the physical character of the stimulus change itself. To ask for a definition of reinforcers which will tell us beforehand for a given species which objects or stimuli will exert the reinforcing effect is to ask that a definition should tell us what the world is like before we investigate it, which is not possible in any science. It happens that the psychologist is worse off than others, because species differences, individual hereditary differences, and differences of the reactional biography make a larger mass of facts necessary in order to know whether a given agent will reinforce a particular organism. But at worst the Weak Law of Effect in conjunction with its member laws is far from

useless. When I know inductively that all non-toxic substances containing sugar will act as reinforcers for organisms from rat to man and therefore that I can almost certainly strengthen all responses learnable by any of these species on the basis of the presentation of any of these substances, I know a great deal and my science has a very considerable predictive power.

AN ANALOGOUS PROBLEM IN PHYSICS

It is instructive to consider a somewhat analogous problem in physics, in the definition of "force." Once mass has been defined by some such artifice as Mach's acceleration-ratio technique, and acceleration defined in terms of time and distance, Newton's second law is a *definition* of force. I neglect here other attempts to introduce the notion such as the "school of the thread" (18), utilizing Hooke's law in the form of a definition rather than a law, or its modern variants, e.g., Keenan's (13) recent effort. Force is "that which accelerates mass." Mach's introduction of the concept of mass was somewhat disturbing to certain of his contemporaries because of a suggested circularity. Mach saw that it was the *inertial* character of mass, rather than "weight" or "quantity of matter" which was crucial in setting up the definition of force. Accordingly, he proceeds as follows:

a. *Experimental Proposition.* Bodies set opposite each other induce in each other, under certain circumstances to be specified by experimental physics, contrary *accelerations* in the direction of their line of junction. (The principle of inertia is included in this.)

b. *Definition.* The mass-ratio of any two bodies is the negative inverse ratio of the mutually induced accelerations of those bodies.

c. *Experimental Proposition.* The mass-ratios of bodies are independent of the character of the physical states (of the bodies) that condition the mutual accelerations produced, be those states electrical, magnetic, or what not; and they remain, moreover, the same, whether they are mediately or immediately arrived at.

d. *Experimental Proposition.* The accelerations which any number of bodies A, B, C. . . induce in a body K, are independent of each other. (The principle of the parallelogram of forces follows immediately from this.)

e. *Definition.* Moving force is the product of the mass-value of a body into the acceleration induced in that body. Then the remaining arbitrary definitions of the algebraical expressions "momentum," "vis viva," and the like, might follow. But these are by no means indispensable. The propositions above set forth satisfy the requirements of simplicity and parsimony which on economic-scientific grounds, must be exacted of them. They are, moreover, obvious and clear; for no doubt can exist with respect to any one of them either concerning its meaning or its source; and we always know whether it asserts an experience or an arbitrary convention (17, pp. 243-244).

In the appendix to the second English edition, Mach replies to critics of this procedure as follows:

A special difficulty seems to be still found in accepting my definition of mass. Streintz has remarked in criticism of it that it is based solely upon gravity, although this was expressly excluded in my first formulation of the definition (1868). Nevertheless, this criticism is again and again put forward, and quite recently even by Volkmann. My definition simply takes note of the fact that bodies in mutual relationship, whether it be that of action at a distance, so called, or whether rigid or elastic connexions be considered, determine in one another changes of velocity (accelerations). More than this, one does not need to know in order to be able to form a definition with perfect assurance and without the fear of building on sand. It is not correct as Höfler asserts, that this definition tacitly assumes *one and the same force* acting on both masses. It does not assume even the notion of force, since the latter is built up subsequently upon the notion of mass, and gives then the principle of action and reaction quite independently and without falling into Newton's logical error. In this arrangement one concept is not misplaced and made to rest on another which threatens to give way under it (17, pp. 558-559).

It is obvious that Mach defines mass in the way he does *so that* the definition of force by $F=ma$ will lead to the kinds of laws we want. That is, a previous "knowledge" of the law of gravity based upon a cruder notion of mass is involved historically in the formulation of such a definition. But the crucial point is that it is involved only in the context of discovery, not in the context of justification (21, pp. 6-7). There is nothing wrong with making use of any notions, including vague anthropomorphic experiences of pleasure, in deciding how we shall formulate definitions, since our aim is to erect concepts and constructs which will fit into the most convenient and powerful system of laws. The point is that we wish to come out with explicit notions that are free of this vagueness and which do not require any notions which cannot be finally introduced objectively. There is probably a remnant of hedonism in the thinking of the most sophisticated contemporary reinforcement theorists, and there is no reason why anybody should pretend that when he talks about rewards he does not have some faint component in his thinking which involves the projection of such pleasure-pain experiences. But this does not mean that these notions are made part of the scientific structure he erects, in the sense that either the definitions of terms or the establishment of the laws requires such associated imagery in his readers. I suggest that Thorndike's critics are in the same position as Mach's.

One might ask, why would a physicist be upset should he attend a spiritualist seance and find tumblers leaping off tables and floating through the air? If the concept of force is given simply by the relation

$F=ma$, then, if a glass tumbler undergoes an acceleration, a force must act and his definition assures him that the physical world will not surprise him. I do not think the answer to this question is far to seek. While it is admittedly a question of decision, I doubt that most physicists would decide to say that an acceleration occurred in the absence of a force. If the genuineness of the phenomenon were satisfactorily established, I do not think there would be a re-definition of the *concept* of force, but rather that the existence of "forces" on other bases than those previously known would be assumed. That is, the physicist would not say "here is a case of acceleration without a force," but he would rather say "here is a case of force not arising from the usual mechanical, gravitational, or electro-magnetic situations which I have thought, up to now, were the sole bases on which forces came into being." It is certainly no criticism of a Newtonian definition of force (I leave out the fact that Newton, while he defined force in this way, apparently also treated his second law as one of empirical content) to say that having thus defined force you cannot know beforehand what are the conditions in the world under which forces will appear. The mechanical forces involved in direct contact, the force of gravity, and certain electrostatic and magnetic forces were known to Newton. There is nothing about his definition of force which tells us that a peculiarly directed force will exist between a wire carrying an electric current and a compass needle, nor that attracting or repelling forces will exist between parallel wires each of which carries a current. The discovery of these conditions under which forces exist was an empirical contribution of Oersted and Ampere.

Similarly, the psychologist defines what is meant by a reinforcer, and proceeds to search for the agents that fall under this definition. There are undoubtedly kinds of stimulus changes of which we are as yet unaware which will turn out to have the reinforcing property. Dr. Wilse Webb (personal communication) has found in preliminary experiments that at least in one kind of Skinner box the click produced by the operation of an empty magazine will exert a reinforcing effect in an animal whose experience has never given this stimulus an opportunity to acquire secondary reinforcing properties. This is surprising to us. What are the conditions under which this will occur? Suppose it should be found that almost *any* stimulus change within a fairly wide range (avoiding extreme intensities which are anxiety-producing) would exert a slight reinforcing effect in the Skinner box or in any similar apparatus in which there is a considerable stimulus restriction and a marked constancy in the homogeneity of the visual and auditory fields. It might be discovered that when a member of this species has remained in

such a homogeneous field for a period of time, stimulus *changes* (not otherwise specified) exert a reinforcing effect. Maybe the rat is "bored" and just likes to make something happen! A difficult notion to nail down experimentally, to be sure. But its complexity and the number of things to be ruled out, does not take it out of the realm of the confirmable.

Let us consider a very extreme case. Suppose in the T-maze situation a systematic increase in the strength of the right turn should be discovered for a standard rat. Suppose that the most thoroughgoing, exhaustive manipulation of the external effect of right-turning should fail to reveal any condition necessary for the effect. "No member of the reinforcing class is to be found." I think that at this point we would begin reluctantly to consider a reinforcing property of the response itself. Perhaps turning to the right is inherently reinforcing to this species. It seems, for instance, that "fetching" behavior in certain species of dogs is self-reinforcing (or at least that it has a biologically replenished reserve). The only reason for calling right-turning "self-reinforcing" rather than simply saying that it is a response of innately high strength in the species is that a *change* in strength occurs with successive runs, otherwise "turning to the right" is simply a kind of tropism. Is the "self-reinforcing" idea factually empty? Although many people would disagree with me at this point, I do not think it is. But it has factual meaning only intradermally. There is no reason why we could not study the proprioceptive effects of a right turn and find out whether, if they are cut out, the increase in response strength continues to occur. In principle we could create the proprioceptive effects of a right turn by artificial means and on that basis strengthen a topographically different response such as lifting the fore paw, wiggling the whiskers, or the like. Here there are difficulties, but I would be prepared to argue that in principle the self-reinforcing effect of right-turning is an empirically meaningful notion.

An interesting side-light is that even the Strong Law of Effect is, as stated, compatible with the latent learning experiments. I am not interested in avoiding the consequences of those experiments by shrewd dialectics, but in the interests of clarity it should be pointed out that in, e.g., the Blodgett design, the big drop in errors *does* follow a reinforcement. So long as the Strong Law of Effect is stated qualitatively and does not explicitly mention amounts and times, it would be admittedly difficult to design an experiment in which it could be refuted. A neo-Hullian interested for some reason in preserving the Strong Law of Effect might simply add a quantitative postulate. He might assume

that when a response undergoes an increment in strength on the basis of a minimally reinforcing agent (that is, one in which the asymptote of the acquisition of habit strength is relatively low), then, if subsequently a strong reinforcement is introduced, the parameter in the new growth function which determines the rate of approach to the new asymptote is greater than it would have been without the original learning. Since in the Blodgett design there is evidence of acquisition of differential habit strengths during the latent phase, such a postulate would lead to a preservation of the Strong Law of Effect. The main reason that we are concerned to deal with latent learning material of the Blodgett type is that in the reinforcement theory as now formulated, the effect of a reinforcer is implicitly assumed to operate immediately.

RELATIONSHIP OF REINFORCEMENT TO DRIVE

Perhaps a comment is needed on the way in which reinforcement has been treated here as the primary notion whereas drive, need, or demand is defined in terms of it. I do not mean to imply that need or drive is not the more "basic" factor, if by this is meant that what is a reinforcer or what acquires reinforcing properties depends upon a certain relevance to need. But this manner of speaking refers to the causal reconstruction of behavior, and reverses the epistemological order. The needs of an organism are inferred from changes in behavior strength as a function of certain states of affairs. That is to say, we "get a fix" on a need by being able to induce the chief defining properties of those states of affairs to which behavior is shown to tend. I do not see how there is any possibility in proceeding otherwise at the level of molar behavior. Whether it will be feasible or desirable to hypothesize a kind of state called need in the case of all reinforcers is a moot point at present. I gather that Hull would argue it will, whereas Skinner would argue it will not. One can consider a sort of continuum of reinforcing states of affairs at one end of which it is most easy and natural and obviously very useful to speak in terms of a need, e.g., the case of food or water; whereas at the other end, e.g., the reinforcing effects of hearing a click or turning off a light, the notion of needs seems relatively less appropriate. But the *causal* primacy of needs in our final reconstruction of behavior laws must not be confused with the epistemological status of needs, i.e., the operations by which we arrive at a conception of the needs. Whether the reduction of need is a necessary condition for learning is a question that is not involved in my formulation of either the Weak or the Strong Law of Effect since need-reduction is not equated to reinforcement. This independence of the notions of rein-

forcement and need-reduction is seen not only in the question of whether need-reduction is (for a sophisticated organism) a necessary condition for reinforcing effect, but it is the intention of these definitions to leave it an open question as to whether a kind of event called need-reduction is involved in reinforcing effects at any stage. The alternative to this is to exhaust completely the concept of need by defining an intervening variable via a class of reinforcing agents, i.e., the organism's "need" is not specified in any way except to say that it is "whatever state" within the organism is involved in the reinforcing effect of a stimulus change known experimentally to exert such an effect. In this case, of course, a rat may be said to have a "need" to keep the light off, to be with another rat, to hear a sound, etc. Whether this is a desirable way of speaking we need not consider here.

In the preceding developments, I have avoided consideration of refinements which would be necessary to complete the theoretical picture. The most important of these is the apparent exception to the Weak Law of Effect in which a change in strength does not occur in spite of the presentation of a known reinforcing agent because certain other dominant factors are at work. As an example, we may consider the "fixation" of a response which is followed by anxiety reduction to the point that an opposing response consistently reinforced with food fails to develop an increase in strength. In any particular situation it is the task of experimental analysis to show what the relations are; as a nice example of this I may refer to the recent work of Farber (3). Of course, if the response does not have sufficient opportunity to *occur*, be reinforced, and hence develop strength, the Weak Law of Effect is not violated. Those cases in which this is not an adequate explanation must be dealt with by considering the opposing forces, leaving open the question as to whether these opposing forces can themselves be satisfactorily subsumed under the Strong Law of Effect. The case here is not essentially different from the case in mechanics where we introduce the concept of force as a dynamic concept (that is, by accelerations produced) and subsequently apply the same notions to systems which are in equilibrium. In physics, one makes use of the laws about force which are based upon the dynamical notion of it in order to explain those cases in statics in which no motion results. Whereas the detailed reconstruction of the causal system remains as a task for the future, I do not believe there are any fundamental logical difficulties involved in the notion that a reinforcing state of affairs is initially defined by an increase in strength, and subsequently the failure of such a state of affairs to exert the effect is explained in terms of the occurrence of other operations or states which oppose it.

SUMMARY

Let me conclude by summarizing the development, using Mach as a model. For convenience I neglect here the specification of a schedule:

a. *Experimental Proposition*: In the rat, if turning to the right in the T-maze is followed by the presentation of sunflower seeds, the strength of the right-turning response will increase. (A situational-reinforcer law.)

b. *Definition*: A stimulus or stimulus change which occurs as the strengthening condition in a situational-reinforcer law is a *reinforcer*.

This empirical law together with the above definition enables us now to assert (as an empirical statement) "sunflower seeds are a reinforcer." The empirical content of this is that there is at least one response which the presentation of sunflower seeds will strengthen.

The presentation of a reinforcer is called *reinforcement*.

c. *Definition*: If the strength of a response may be increased as a function of behavior in an exposure to a situation (rather than by surgical, drug, or maturational changes), such a response is *learnable* by the organism. No reference to reinforcement is made here; we simply require that response strength be shown to increase following "experience," of whatever sort.

d. *Experimental Propositions*: Following suitable manipulation of their experiences, rats will show increases in the strength of pressing levers, pulling chains, rolling marbles, turning to the right at certain choice points, gnawing through paper, digging through sawdust, turning wheels, etc. (Expanded, this would consist simply in a long list of specific "laws" asserting the learnability of certain response classes.)

e. *Experimental Propositions*: Sunflower seeds may be used to strengthen lever pressing, chain pulling, etc. In general, sunflower seeds may be used to strengthen all learnable responses in the rat. (This asserts the generality of the reinforcing effect of sunflower seeds and is what I am calling a trans-situational reinforcer law.)

f. *Definition*: A trans-situational reinforcer is a stimulus which will strengthen all learnable responses. (We have already defined reinforcer so that it does not commit us to its generality, that is, a reinforcer is *at least* a situational reinforcer. If there are any reinforcers which exert the reinforcing effect upon all learnable responses, they are trans-situational). This definition with the immediately preceding experimental propositions enables us to say, "Sunflower seeds are a trans-situational reinforcer."

Such a collection of specific empirical laws in combination with the above general definition leads to a large set of laws such as these last stated ones so that in the end we find the following:

g. *Experimental Proposition*: All reinforcers are trans-situational. (The Weak Law of Effect.)

h. *Experimental Proposition*: Every increment in strength involves a trans-situational reinforcer. (The Strong Law of Effect.)

It seems clear that in the above sequence both the definitional and the factual (empirical) elements are present, and in a simple, commonplace form. The definitional and conventional elements appear in the specification of the circumstances under which a stimulus is to be called "reinforcing." Such a stipulation, however, cannot tell us whether any such stimuli exist. That they do exist, which no one doubts, is an empirical finding; and the numerous statements about them constitute situational-reinforcer laws which are in a sense the special "sub-laws" of effect. These are related to the Weak Law of Effect somewhat in the same way that the particular empirical laws about the properties of bromine, fluorine, chlorine, and so on, are related to the Periodic Law. That the stimuli which occur in the situational-reinforcer laws have a generality of their reinforcing power is also an empirical finding, at present less well established (the Weak Law of Effect). That all cases of learning require certain time relationships to the presentation of such general reinforcers is yet a further factual claim, at present very much in dispute (the Strong Law of Effect).

I can see no reason why any theorist, whatever his position, should find the preceding treatment objectionable as an explication of the Law of Effect. I do not see any way in which the Strong Law of Effect, which is after all the big contemporary issue, has been surreptitiously put into the definitions in such a way that what is intended as an empirical proposition is effectively made a consequence of our use of words. The status of the Strong Law of Effect and even to some extent the Weak Law is presently in doubt. Further, some of the words used in these definitions, e.g., the word "response," are difficult to define in a way that makes them behave in the total system as we wish them to. I have not tried to deal with all these problems at once, but I hope that there are no difficulties springing from the problem of circularity which have not been met. That it is difficult to untangle the learning sequence which has given the reinforcing property to some states of affairs, particularly in the human organism, is admitted by everyone. That a large amount of detailed work of the "botanizing" type, not particularly ego-rewarding, needs to be done before the special sub-laws of effect are stated in terms of quantitative relations is quite clear. Finally, it would be very nice if in some magical way we could *know* before studying a given species exactly what stimulus changes would have the reinforcing property; but I have tried to indicate that this is an essentially irrational demand. In the light of the previous analysis I think the burden of proof is upon those who look upon a sophisticated formulation of the Law of Effect as circular, in either of the ordinary uses of that word.

BIBLIOGRAPHY

1. CARR, H. A., *et al.* The Law of Effect: a roundtable discussion. *Psychol. Rev.*, 1938, **45**, 191-218.
2. ESTES, W. K. A study of motivating conditions necessary for secondary reinforcement. *Amer. Psychologist*, 1948, **3**, 240-241. (Abstract.)
3. FARBER, I. E. Response fixation under anxiety and non-anxiety conditions. *J. exp. Psychol.*, 1948, **38**, 111-131.
4. FEIGL, H. Operationism and scientific method. *Psychol. Rev.*, 1945, **52**, 250-259.
5. FEIGL, H. Some remarks on the meaning of scientific explanation. In H. Feigl & W. Sellars., *Readings in philosophical analysis*. New York: Appleton-Century-Crofts, 1949. Pp. 510-514.
6. GUTTMAN, N. On the relationship between resistance to extinction of a bar-pressing response and concentration of reinforcing agent. Paper presented at the meeting of the Midwestern Psychological Association, Chicago, Ill., April 29, 1949.
7. HERON, W. T., & PEAKE, E. Qualitative food deficiency as a drive in a discrimination problem. *J. comp. physiol. Psychol.*, 1949, **42**, 143-147.
8. HILGARD, E. R. *Theories of learning*. New York: Appleton-Century-Crofts, 1948.
9. HILGARD, E. R., & MARQUIS, D. G. *Conditioning and learning*. New York: Appleton-Century, 1940.
10. HOSPERS, J. On explanation. *J. Philos.*, 1946, **43**, 337-356.
11. HULL, C. L. Thorndike's *Fundamentals of learning*. *Psychol. Bull.*, 1935, **32**, 807-823.
12. HULL, C. L. *Principles of behavior*. New York: Appleton-Century, 1943.
13. KEENAN, J. Definitions and principles of dynamics. *Sci. Mon.*, N. Y., 1948, **67**, 406-414.
14. LENZEN, V. F. *The nature of physical theory*. New York: John Wiley, 1931.
15. LINDSAY, R. B., & MARGENAU, H. *Foundations of physics*. New York: John Wiley, 1936.
16. MACCORQUODALE, K., & MEEHL, P. E. "Cognitive" learning in the absence of competition of incentives. *J. comp. physiol. Psychol.*, 1949, **42**, 383-390.
17. MACH, E. *The science of mechanics* (Transl. by T. J. McCormack). Second English Ed. Chicago: Open Court Publishing Co., 1902.
18. POINCARÉ, H. *The foundations of science*. New York: Science Press, 1913.
19. POSTMAN, L. The history and present status of the Law of Effect. *Psychol. Bull.*, 1947, **44**, 489-563.
20. PRATT, C. C. Operationism in psychology. *Psychol. Rev.*, 1945, **52**, 262-269.
21. REICHENBACH, H. *Experience and prediction*. Chicago: Univ. of Chicago Press, 1938.
22. SKINNER, B. F. *The behavior of organisms*. New York: Appleton-Century, 1938.
23. SPENCE, K. W. Studies on secondary reinforcement. Address given to the Minnesota Chapter of Psi Chi, Minneapolis, April 22, 1948.
24. TAYLOR, L. W. *Physics, the pioneer science*. New York: Houghton, Mifflin, 1941.
25. THORNDIKE, E. L. *The fundamentals of learning*. New York: Teachers College, Columbia Univ., 1932.
26. THORNDIKE, E. L. *Animal intelligence*. New York: Macmillan, 1911.
27. THORNDIKE, E. L. *The original nature of man*. New York: Teachers College, 1913.
28. THORNDIKE, E. L. *The psychology of learning*. New York: Teachers College, 1913.
29. TOLMAN, E. C. *Purposive behavior in animals and men*. New York: Appleton-Century, 1932.

Received July 27, 1949.

BOOK REVIEWS

THORNDIKE, ROBERT L. *Personnel selection test and measurement techniques*. New York: John Wiley, 1949. Pp. viii+358. \$4.00.

In one sense, the health of a science may be assessed upon the basis of its ability to advance when granted unusual opportunity, and the manner in which that growth is reflected in its texts and reference volumes. Psychologists rose to the challenge and opportunity given them in the field of selection by the war and demonstrated that they could perform a useful function while increasing their fund of knowledge and understanding. Now Thorndike has shown that the results of this work can be brought to the student or layman in a well-organized, concise, and internally consistent account of basic principles and techniques.

Indeed, this book fills a need which has been rapidly approaching the acute. The confusion and downright misinformation which mark the discussion of selection testing in most of our textbooks have done much to encourage the tendency toward test validation by analogy or fiat. This tendency is now appearing in the work of industrial, clinical, and guidance psychologists alike. The author's powerful presentation of basic theory, accompanied by his consistent acceptance of the demands of logic and scientific method, should do much to clear the air of the fallacious and the opportunistic.

The first section of the book is devoted to the techniques of building and evaluating a test selection program. Beginning with a chapter in which job analysis is properly recognized as ". . . a source of hypotheses for selection tests and a source of insight about criterion measures" (p. 13), Thorndike proceeds to a discussion of various types of selection instruments and their construction. This discussion is sound and comprehensive. It is in the following chapters on reliability and validity, however, that the author makes new and important contributions. The discussion of reliability is pursued with originality and a firm refusal to oversimplify what is really a complex and vitally important concept. The conclusion that "It becomes clear that the basic problem for determining the reliability of any measurement procedure becomes that of *defining* what shall be considered true variance between individuals and what shall be thought of as error variance" (p. 72), and the development which follows should serve to make many who believed they had mastered the concept approach it with a new respect.

The chapters on validity are of equal stature. The problem of criteria, shamefully neglected for so long by so many, at last receives a precise and careful analysis. Not the least of the author's virtues is his ability to unearth and illuminate unsolved or difficult problems. When the reader has finished following Thorndike's consideration of criteria, he cannot fail to recognize at one time how stringent are the require-

ments for satisfactory performance measures and with what difficulty they are met. On the subject of the estimation of test validity, there is, once again, a skillful dissection of many of the complexities which have been skirted in most texts. Having made a point which is only too often ignored in the interests of convenience, "Validity coefficients based on fewer than 100 cases are too unstable to serve as the basis for any serious comparative study of different tests for personnel selection . . ." (pp. 58-59), the author considers validity estimation under various combinations of criterion and test score with regard to the factor of continuity or discreteness. The assumptions underlying the use of various correlational measures are clearly stated and the pitfalls encountered in their interpretation receive the appropriate danger signals. Of particular interest is the treatment of population curtailment—an ubiquitous but neglected difficulty.

Chapters on the combination of tests into a battery and the analysis of test items include an evaluation of the clinical method of test combination and interpretation. The author's restrained but firm insistence upon the fact that the clinical method is, of its nature, specific to the clinician and the personnel situation in comparison to the generality of the objective method of linear combination stands in clear contrast to the fuzzy thinking which has resulted in attempts to distinguish between empirical validity and "clinical validity," whatever the latter term may mean.

The author then proceeds to a very practical consideration of the administration of testing programs. This account, covering such topics as the conduct of testing, scoring and weighting procedures, the organization of reports and records, quotas and yields, and the effective public presentation of test research and results, should serve two important functions. For the student, these chapters should help to bring recognition that personnel testing is not performed in a vacuum but is, instead, applied to real people doing everyday jobs. For the personnel executive and for the psychologist whom he employs, the chapters should produce a respect for the necessity to understand what functions the testing program is to perform and how it may be designed to fulfill them best. In this connection, the author's discussion of the relationship between selection, multiple selection, and classification is particularly useful.

Of course, the book has some faults. It is probable that the author would be the first to admit that he has not solved the ever-present difficulty of insuring that the reader will understand the statistical concepts so necessary to a grasp of this material. His implication that a background corresponding to about a year of work in statistics is sufficient for his readers seems unrealistic to this reviewer. Students will need help and plenty of it in most of the chapters. The writing is clear and readable, but there has been no attempt to predigest or to prettify.

The reader is expected to work and so, quite properly, is the instructor.

The wide adoption of this work as a text in courses on personnel seems inevitable. If those who test and interpret also adopt the integrity of the book's approach and its basic philosophy, psychology will be considerably in Thorndike's debt.

S. RAINS WALLACE, JR.

Life Insurance Agency Management Assn.

CHASE, STUART. *The proper study of mankind. An inquiry into the science of human relations.* New York: Harper, 1948. Pp. xx+311.

The books of Stuart Chase are neither popularizations nor original contributions to knowledge and science. Chase is a journalist of highest character and integrity in the fields of social science. In earlier books he has reported on subjects ranging from the waste of natural resources to semantics. The purpose of his latest book, he tells us, is "to run a kind of chain and compass line across the whole front of the sciences devoted to human relations." This is a formidable undertaking. The result does not wholly satisfy Mr. Chase nor is it likely to satisfy the specialist in any of the fields involved. But we may be grateful for the attempt.

The book is organized into three parts. Part I develops the conception of science as it is generally accepted today and illustrates the contention that social relations are subject to investigation by the scientific method. The point of view has been well developed before, but this is as clear and simple a statement of it as this reviewer has read. It is particularly effective in making the point that social science already has a small but significant body of empirically established generalizations "in the storehouse" which have been useful, and can be further useful, in application to problems of social relations. Chase debunks the idea that scientists are magicians suddenly producing an atomic bomb or a "grand solution." But the air of breathless suspense with which some of the researches are reported presents an equally unrealistic description of the painstaking little steps by which a science is built.

Part II, entitled Landmarks and Achievements, presents selected researches and key ideas primarily developed out of the cross-fertilizations of sociology, social anthropology, and social psychology. It also presents some comments on the dismal shortcomings of economics and political science as sciences. There are chapters summarizing Leighton's report on Poston, the current ideas on race and inter-group relations, Ogburn's theory of cultural lag, the social anthropologists' work on American communities, the Western Electric studies, and public opinion polls. These, and some others, are discussed with skill but with the inevitable over-simplification required to make them "landmarks." The integration of the section suffers from the fact that such researches as Kinsey's on the sexual behavior of American males and the strike in the shoe factories of Yankee City are not parts of an integrated theory of social behavior. This is not Mr. Chase's fault.

In Part III Chase attempts to pull the results of his summary together and point up the moral of his inquiry. Here he discusses communication as a basic problem with a separate chapter on the contributions of semantics. He cautions reformers to reform if they are to lay the scientific foundations for achieving one world. He points the "roads" to answers to practical problems: a gigantic project to plan for permanent peace, more teamwork, encouragement of social science through better recruitment of talented young people, more funds, and more public awareness.

The practical problems offered in this section as ultimate objectives of social science are usually stated in economic or political terms. But economics and political science are the very fields which least fulfill Mr. Chase's conception of science. This is partly because these fields lack the underpinning of the other social sciences and partly because they fail to use what is available to them. The basic reason is, however, that most economists and political scientists ask questions not of science but of policy.

Political science and economics are the senior members of the social science family but they have been the last to turn away from the problems of what *should* be. Perhaps their older emphasis should continue. Understanding of behavior—including economic and political behavior—can be approached through the sciences of human relations, but the directions of public policy must be derived from other sources. There is a place—possibly within a broadened conception of social science—for the disinterested consideration of the ends of social policy to which the instruments and findings of science can be put. Chase might well have considered this orientation as an alternative one for economics and political science.

Mr. Chase is not equally critical of all the social sciences. The shining examples of research useful in practical situations seem partially to blind him to the confusions of terminology and the shortcomings of systematic theory in social psychology, sociology and anthropology. Perhaps the general public to whom the book is primarily addressed needs such over-statements tempered only by the qualifications which Mr. Chase includes. Specialists in these fields will be inclined to think that their sciences have almost been over-sold. Despite the reminders from time to time that social science is in the early stages of its career, the net impression which the book is likely to give to the uninitiated reader is that social science is ready for application to many problems today. Actually it can make only limited contributions. It is no service to social science to build up expectations for its performance which it can only partially fulfill.

The book has been offered by the publisher as a textbook for college courses. In this role it falls short because of the very characteristics which make it a good book for the trade. It is neither systematic nor detailed enough to serve as an adequate summary of the sciences of

human relations. It will, however, make excellent collateral reading. No other book has come to the reviewer's attention which gathers together in one place such lucid reports of good current research in all the major social sciences. For this reason, if for no other, it will serve a useful function for the departmentalized college student.

HENRY J. MEYER.

New York University.

SMITH, GUDMUND. *Psychological studies in twin differences*. Lund, Sweden: Gleerup, 1949. Pp. 251.

This is essentially a monograph reporting the author's investigations on 61 pairs of like-sex twins ranging in age from 8 to 69 years and living in the vicinity of Lund, Sweden. The twins were classified as identical or fraternal on the basis of birth records as well as a series of anthropometric characteristics including fingerprints, blood groups, iris color and structure, hair and skin coloring, and facial and cranial measurements. The total group contained 32 identical pairs (13 male and 19 female) and 29 fraternal pairs (16 male and 13 female); these numbers include one set of triplets who were entered as one identical and two fraternal pairs in most of the calculations.

Two parallel approaches were followed. One included the administration of imagery tests, including measures of duration, periodicity, and size of a negative after-image when projected at different distances, and an intensive study of eidetic imagery. Evidence of the latter was found in 25% of the subjects under 16 years of age and 9% of the older subjects. The second approach comprised a personality study, including two questionnaires answered by the twins themselves and two by their families, supplemented in some of the cases by teachers' or employers' ratings, school records, and a "pair test" in which a difficult construction puzzle was attempted first by each twin separately and then by the two working together, their reactions under both conditions being observed by an assistant. The personality data and the more qualitative results of the imagery tests are reported principally in the form of case studies of each pair, while the quantitative data are treated in terms of analysis of variance and intra-class correlation. It should be noted that the number of twin pairs on whom these various techniques were employed varied; the imagery tests, which constituted the main part of the study, were given to only 22 identical and 26 fraternal pairs.

Throughout the study, special attention was given to intra-pair relations and the specialization of roles. The relations of identical twins seemed to be characterized more often by coordination and cooperation, those of fraternal twins by competition and conflict. The influence of such factors as age, sex, intellectual level, and personality characteristics upon twin relationships is also discussed in detail. A number of imagery characteristics yielded significantly higher intra-pair correlations among identical than among fraternal twins. The investigator points out re-

peatedly that the higher concordance of identicals was more clearly evident in certain relative indices than in isolated measures of reaction. For example, the absolute size of the images showed no greater correlation among identical than among fraternal twins, but change in image size with repetition did show closer resemblance among identicals. Another example of these relative measures, which the author terms indices of "development" of the image, is based on the effects of "intermittence" upon size, duration, and periodicity of the image. Such "intermittence" was achieved by inserting a rotating disc with two black sectors in front of the eye-piece, while the subject was projecting the image. The investigator also purports to show that these measures of "development," as contrasted to isolated static scores, are more likely to correlate with basic personality characteristics. For a demonstration of this relationship, the author turns to the case study material on the same subjects.

On the basis of the above findings regarding the possible value of measuring *changes* of response under varying conditions, the author concludes that "every test should be a longitudinal study, however brief" (p. 233). This emphasis is probably the most suggestive contribution of the monograph to general experimental design. With respect to specific findings, two cautions need to be observed in interpreting the data. First, the number of cases on which any one comparison was based was often very small. Secondly, a strong element of subjectivity entered into most of the observations, especially those dealing with broad personality characteristics. Thus in the "pair test," no quantitative records are given, but the observations of the assistant have simply been incorporated into the general case reports. The questionnaires were used primarily as a basis for informal questioning by the experimenter, and again no specific replies are reported. Even more subjective is the subsequent analysis and classification of case reports and the linking of personality patterns with imagery characteristics. The role of experimenter's set cannot be discounted under these conditions. Finally, it might be added that the style and organization of the monograph make it very difficult to read. The prevalence of literal translation produces some extremely awkward passages as well as a very unusual use of English words. For example, the word "formula" as herein used signifies "questionnaire" and case studies are described as "casuistics." The entire presentation could have been considerably shortened and clarified.

ANNE ANASTASI.

Graduate School, Fordham University.

JOHNSON, DONALD M. *Essentials of psychology*. New York: McGraw-Hill, 1948. Pp. xiii+491. \$3.50.

For many years, text-books designed for use in the beginning course in psychology have shown a tendency toward expansion. The author of the present book has reversed this process and produced one of the

shortest texts for the beginning course that has been published in recent years. Of the 491 pages in the book, 442 contain the actual text, the remaining pages being devoted to references, a glossary, a list of audiovisual aids, and index. That the general content of the beginning course is covered in this short space is indicated by the chapter contents: motives and emotions, sensory functions, attention, perception, learning, thought and judgment. A little over half of the book (228 pages) is devoted to the preceding topics. The remaining 214 pages are divided into four chapters on the individual and society, abilities and tests, personality, and abnormal personalities. There is a summary at the end of each chapter and a list of technical terms for study. The book contains a total of 122 figures relevant to the text discussion. In general, the style is readable, concise, and informal.

The guiding principles which have aided the author in producing a text as short as this are described in the Preface. The importance of integration is stressed and principles which are referred to later in the text are therefore placed in the earlier chapters. Historical material and discussions of psychological schools are omitted. No attempt is made to document and prove every point. Classifications (of methods, behavior, etc.) are omitted unless they are used at least twice in later exposition. The author's hope that a book guided by these rules would be shorter than usual has been realized.

With a text as short as the present one, the obvious question arises: Is it possible to present the essentials of psychology adequately in 442 pages? The answer to this question will undoubtedly vary with the individual who reads the book. It is possible to find sections or topics which one feels to be too scantily treated (e.g., maturation and development). It is possible to find topics which are so concisely treated that they will probably need considerable supplementation before they can be adequately understood by the student. For example, the following constitutes the entire explanation of negative after-sensations: "The negative afterimage can best be understood by remembering that nerves are living tissue and that the chemistry of the receptors tends to remain in equilibrium. When stimulation pushes the photochemical balance in the eye off equilibrium, it tends to return, and the returning is a stimulus for the complementary colors" (p. 83). It is possible for the reader to find some scantily treated topics which he believes to be of greater importance than other topics which are treated at greater length. For example, six kinds of learning (motor skills, modification of perceptions, etc.) are discussed for 14 pages while efficiency of learning receives but three pages. A poem called "Similar Cases" is given a full page (p. 185), while the distribution of practice in learning receives a short paragraph of seven lines. Some instructors may doubt that enough experiments, especially those illustrating important points of methodology, have been

presented in sufficient detail to give the beginning student the flavor of scientific research in psychology.

But in spite of the criticisms that may be made and the differences in opinion that may exist, the author has undoubtedly included much of importance and has presented many of the essentials of psychology in a concise way. He has not merely omitted difficult topics to make an easier and shorter text. Elementary statistical concepts and methods, for example, are presented in an appropriate place. The earlier chapters appear to be the most concise and would probably need the greatest amount of supplementation; the later chapters are longer and more completely developed. One may still feel, as does the reviewer, that the book might be better if it were longer and if more space had been used in discussing some topics. But for those who want a shorter text, *Essentials of Psychology* is well worth considering.

EDWARD E. ANDERSON.

Wilson College.

V. SCHILLER, PAUL. *Aufgabe der Psychologie: eine Geschichte ihrer Probleme*. Vienna: Springer-Verlag, 1948. Pp. 233.

The author, formerly of Budapest and recently a visiting lecturer at the University of Chicago, is now professor at the Bolyai University in Klausenburg, Romania. He gives a scholarly and authoritative survey of the theoretical positions of the major figures in the history of psychology from Plato through Tolman. Especially interesting is his discussion of contradictions which follow from Wundt's attempt to distinguish mediate and immediate experience and of Wundt's failure to follow his own theoretical program, in that he studied human functions and activities rather than conscious experiences. Some of the many other sections which deserve mention are those dealing with the theories of von Uexküll, American functionalism and behaviorism, and Brentano.

The book culminates in the author's *Handlungslehre*. The proper task of psychology is to study actions and their motivation. An action is defined as a specific behavior which is dependent on the total situation at a given time and which changes the situation; it includes the processes of arousal, steering, and execution united in a single whole. Conscious experiences are to be viewed as components of actions and not as belonging to an independent world of their own. The author regards his theories, which are expressed at a very abstract level of discourse, as representing a return to the holistic point of view characteristic of Aristotle and other thinkers who did not attempt to restrict psychology to the study of experiences or of external behavior and as a remedy for the tendencies which threaten to split modern psychology into many diverse and seemingly unrelated branches.

FRANK A. PATTIE.

University of Kentucky.

MACHOVER, KAREN. *Personality projection in the drawing of the human figure*. Springfield, Ill.: Charles C Thomas, 1949. Pp. ix+181. \$3.50.

It would be unfortunate if we came to feel that our techniques in personality study were apart from our theories. If theories are valid, they should be demonstrable. If techniques are valid, it must be because they stem from some hypothesis which is basically true. This should hold for all aspects of our research even when practical considerations are most important, as it is understandable that they should be for the clinician whose primary concern is for the patient's welfare. It is, therefore, to be deplored that a prominent clinician, describing an important projective device, should state that: "Incentive for, and primary focus of, investigation centered around perfection of the drawing technique as a clinical tool for personality analysis, rather than around any theoretical hypotheses" (p. 20).

This statement is, of course, not adequately descriptive of the content of the book. The first part, "Personality Projection in the Drawing of the Human Figure," actually has a section on theoretical considerations. The second part, "Principles of Interpretation," draws heavily on psychoanalytic theory, as well as on empirical clinical experience. The third part, "Illustrative Case Studies," is an attempt to relate the theory of projection in drawing the human figure to examples of drawings accompanied by case histories.

The main purpose of this book, as stated by the author, is to begin to systematize the analysis of personality based on drawings of the human figure. Dr. Machover devotes several pages in the first section to a consideration of the sources of projection, which she labels the "psychic datum." She finds three major ones: common social meanings, the individual's own experience and symbol values. In discussing the common social meanings she writes: "It suffices to say that, from the empirical standpoint, such graphic communications occur regardless of age, skill or culture. One source to consider is the common social meanings that physical attributes tend to acquire in the course of social expression and intercourse" (p. 7). If the meanings are acquired as a consequence of social intercourse, how can they be free of culture, or even of age? Experience with the Rorschach, to which she often refers for corroborative evidence, indicates differences in interpretation for different cultures and for different age groups in our own.

In spite of such theoretical problems, the method seems to have considerable validity in clinical procedure. As one reads the illustrative cases, Dr. Machover's analyses can be followed without much argument. The question which often arises in connection with new projective techniques—"Would I see that if I didn't know the case history?"—is answered by her report on "blind" interpretations (pp. 25 to 27).

The drawing of the human figure is an extremely intriguing projective device. It is, admittedly, still in its infancy and communication of method is limited, as for most techniques which emphasize patterns of relationship rather than "scoring." It is very much to be hoped that readers will be stimulated by this book to further study and research. Dr. Machover introduces some challenging experimental problems. In spite of the format of the book, it is a progress report rather than a manual. It would be undesirable, with so many people working on this new method, to accept this report as a manual and mark the problem solved.

LILLIAN WALD KAY.

New York University.

MAYO, ELTON. *Some notes on the psychology of Pierre Janet.* Cambridge: Harvard Univ. Press, 1948. \$2.50.

The author, in his preface, states that "this book is not intended for medical students or for those who have a special interest in the problems of psychiatry." It was written for "a few colleagues whose work demanded close attention to the difficult social, personal, and administrative problems of our time." While it is probably true that a knowledge of personality dynamics would be of service to such persons, it is hard to see in what way Mayo's presentation would be likely to give them valuable insights.

The book does not attempt a complete account of Janet's work, but his chief concepts are discussed sympathetically. Janet defines hysteria as a condition characterized by easy hypnotizability, suggestibility, dissociation, somnambulism, and shifting analgesia and paralyses. Most patients commonly diagnosed as hysterics Janet classifies as obsessives. It is as clear in reading Mayo's account as in reading Janet himself that many characteristics of his particular group of hysterics are the result of tutoring on the part of the psychiatrist. Mayo, in fact, says: "In several of these patients *Janet developed the capacity for separate response until it was possible to conduct two conversations with them simultaneously*" (p. 30).

Morton Prince, with Miss Beauchamp, demonstrated to what lengths this could be carried. It would be rare indeed were a layman to meet such a person in his daily life. Even the student of social administrative problems could manage quite well indefinitely without making use of this classification. Janet himself, according to Mayo, recognizes that "in the modern world the incidence of 'hysteria,' in the strict sense, has notably diminished" (p. 66). It would be difficult to defend Mayo's assertion that this change has occurred as a result of advances in nutrition and education rather than in improvement in diagnosis and therapy.

The obsessives, who include not only the people so classified ordinar-

ily but some who would usually be considered to have hysteria or anxiety states (patients without somnambulism or divided personality), are the more important group according to Janet. Mayo believes them to be "the proper subject of psychopathological study in the strict sense" and the group most likely to concern a student of society. They may be divided into "frightened people" and "destroyers." They are unhappy, guilt-ridden, unable to relate themselves to others. Feeling that they live in an unreal world populated by phantoms, they are preoccupied with false alternatives, rituals, and fears which make it difficult for them to act or make decisions. It is important, says the author, for all students to understand the complexity of attention and the interrelation of action and reflection in order to detect the appearance of obsessive thinking in a social or industrial situation. The appearance of such symptoms will warn him that there is "some irksome constraint or feeling of insecurity imposed upon the individual or group . . ." If he does not discover the source of this, the students will remain unable to understand the social situation he is studying.

While it is possible that, as Mayo says, "there is no real conflict between the observations of Janet and Freud. Indeed they work in different parts of the same field," this book suggests why Janet's work was relatively less fruitful. The rigidity of the classification and emphasis upon description rather than upon intervening "etiological" variables provide the clue.

JAMES G. MILLER.

University of Chicago.

SYMONDS, P. M. *Dynamic psychology*. New York: Appleton-Century-Crofts, 1949. Pp. vii+413. \$3.75.

This book is a revision, prepared for use as an undergraduate text, of the author's earlier book, *The Dynamics of Human Adjustment*. From the previous book, the following subject matter has been deleted: "(1) psychoanalytic interpretations which are least accessible to consciousness of most persons and hence which seem least convincing; (2) dynamics of abnormal states and conditions; and (3) unfamiliar vocabulary" (p. vi). The opening chapter is a more or less conventional introductory discussion of the relationship of psychology to other sciences, values to be derived from the study of psychology, etc. A new note is introduced, however, when the author says that "hundreds, indeed, thousands of reports of interviewing and analytic sessions have been reported and there is a repetition and consistency in them which go far toward overcoming the subjectivity of separate individual reports" (p. 9). While clinical data are rapidly attaining respectability, Symonds is probably among the first text-book writers to acknowledge explicitly its new status. The remaining chapters emerge from the author's essentially

psychoanalytic frame of reference and deal primarily with the various "dynamisms" and the conditions in the environment and within the personality that govern adjustment.

The individual chapters of *Dynamic Psychology* are well organized and relevant points of view and findings are pulled together from a great number of sources. The reader is presented with what amounts to a review of the literature concerning the topic under consideration but reference is made to various writers in an expository style. The text, therefore, does not suffer from the dryness that might be expected in the ordinary "bibliographical" treatment of subject-matter. The over-all unity of the book could be improved. The author devotes only a paragraph to his general theme in the first chapter. In this paragraph he says, "This book deals with *dynamic psychology* which is interested in many states and processes, and particularly with those driving forces within the individual which seek their satisfaction in the world about" (p. 7). While this statement is satisfactory, it seems to the reviewer that the author might have done well to elaborate upon it toward the beginning of the book in his chapter on "The Ego and the Self." This chapter is the next to the last. Since it deals with the processes involved in the individual's relating himself to the environment there would be some advantage to bringing it forward and using it for purposes of developing a unifying theme.

Symonds has managed to write the kind of book that can be adapted to many purposes. The liberal use of illustrations and short cases makes it suitable for courses in adjustment that are primarily concerned with fostering the student's understanding of himself and others. The content can be counted upon to provoke discussion of various aspects of interpersonal relationships. The instructor who prefers a more formal presentation of theories of motivation, adjustment, and "personality" can also take his departure from the author's explanation of concepts, findings, and points of view of various writers. No attempt is made, however, to present a single systematic position. The instructor who is not content with an eclectic approach will not find the definitive framework he is looking for in this book, but he will find that it includes the ideas of a great many people.

FRANKLIN J. SHAW.

Purdue University.

THORPE, L. P., & KATZ, BARNEY. *The psychology of abnormal behavior*. New York: Ronald Press, 1948. Pp. xvi+877.

THORPE, L. P., & KATZ, BARNEY. *Workbook in the psychology of abnormal behavior*. New York: Ronald Press, 1948. Pp. x+124.

This encyclopaedic volume (877 pages) is intended for use as the textbook in the undergraduate introductory course in the psychology of

abnormal behavior. The unusual range of topics covered, the logical and clear organization of the book as a whole and of each chapter, and the concise and direct style in which the book is written make it well suited for its purpose.

The major emphasis is placed on the dynamics of abnormal behaviors, in an effort to lead the student to an understanding of the mechanisms and adjustment processes utilized by maladjusted personalities in their efforts to maintain their psychological integrity. The authors appear to have achieved an unusually eclectic coverage of the dynamic and etiological factors involved in abnormal behaviors by presenting the major competing views.

Part I (39 pages) is a brief but adequate introduction which reviews the social status and significance of abnormal behaviors, the history of research, points of view, and therapies, and the major contributions of foreign psychologists, psychiatrists, and others to the field.

Part II (115 pages) presents a general introduction to the dynamics of abnormal behaviors, criteria used for evaluating behaviors as abnormal, the nature of such behaviors, and a survey of organic and psychological etiological factors. Part III (58 pages) presents discussions of the nature and significance of symptoms, the classification of abnormal behaviors, and a review of abnormalities of the emotional, cognitive, and motor processes and of sleep. These two parts serve as brief but very adequate introduction to fundamental concepts regarding the dynamics, etiology, symptomatology, and evaluation of healthy and unhealthy behaviors. They also serve as adequate prerequisites to the parts which follow.

Parts IV through VII are concerned with specific adjustment problems and maladjusted personalities. The discussion of each disorder is presented in organized form in terms of incidence, symptomatology, etiology, treatment, prognosis, and prevention. This outline may be somewhat stereotyped but it provides the student with a definite study outline.

Part VIII (126 pages) reviews the field of therapies, including discussions of the objectives of therapy, methods of examination, and the major psychoanalytic, psychotherapeutic, and psychiatric methods. The presentation of these methods is primarily descriptive rather than evaluative or critical. The appendices present data on the classification of hospitals, statistics on the frequencies of patients with the various disorders, and a glossary of terms. The latter will be especially helpful to students. The book contains ample illustrative case reports. It is up to date on recent developments in diagnosis and treatment.

The authors have published a workbook for use with the text. It contains an objective test, suggested "research" projects, and discussion or "essay" questions for each chapter in the book. It also contains in table form chapter by chapter references in ten other books in the field

of abnormal behaviors, so that students and instructors are provided with an easily used guide to assignments for collateral readings. Instructors will find the workbook a welcome source of materials for tests, of lists of relevant films and their sources, and of reading assignments.

WILBUR S. GREGORY.

The University of Redlands.

FRANK, L. K. *Projective methods*. Springfield, Ill.: Charles C Thomas, 1948. Pp. vii+86. \$2.75.

BELL, J. E. *Projective techniques*. New York: Longmans, Green, 1948. Pp. xvi+533. \$4.50.

The first book is a little volume which has been written by a man well known to psychologists and other students of human behavior as the director of the Zachry Institute of Human Development in New York City.

The first part of the monograph discusses the growing use of projective methods in psychology and psychiatry and shows how they are predicated upon a psychocultural conception of the emergence of the personality and its dynamic operation. The second part describes recent developments in scientific concepts and methods, especially in physics and chemistry and medicine. The third part reviews various approaches to the study and diagnosis of personality by way of preface to the more specific description of actual procedures. The fourth part outlines projective techniques and describes the five different varieties thus far developed, with interpretative comment upon the value and use of each variety, e.g., constitutive, constructive, interpretative, cathartic, refractive. The fifth part discusses reliability and validity of projective methods and the development of new criteria for assaying projective procedures.

Chapters one to three, approximately 41 pages, are a plea for newer types of psychological testing in light of changing concepts of personality theory. To bolster his position, the author borrows heavily, perhaps too frequently, from events in physics in particular and science in general. In subsequent chapters the author makes many statements without supplying the reference to the literature, which detracts from the book's usefulness to the advanced student. This limitation occurs rather frequently in the discussion of reliability and validity of projective methods.

Frank raises a tremendously important point which research psychologists as well as clinicians must some day attempt to rectify.

While psychologists customarily rely upon the judgment of lay, naive individuals to validate their findings, it is difficult to find any scientific justification for such a procedure. If a scientific method or test is used, one cannot assume that the agreement or disagreement with its findings by a group of untrained, naive judges (as in matching) has any significance. Yet, this reliance

upon such judges is frequently cited as validating or invalidating a method (p. 66).

The appendix contains a partial bibliography of projective methods taken in part from Helen Sargent (pp. 68-86). This is a useful addition but has already appeared in the literature. This little volume is not prepared for the undergraduate major in psychology, nor will it be useful to the graduate student without preliminary training and experience in clinical psychodiagnostic methods. It has merit for theoretical exposition and for indicating the need for more research on projective tests before they earn their place in a scientific psychology.

While Frank's book is a theoretical treatment of projective methods, Bell's volume is also a useful survey of the varieties of techniques, and it attempts to serve more practical purposes. Thus, the author's aim is (1) to present a comprehensive review of the literature on projective techniques so that (2) research workers could apply batteries of projective tests to the same subjects for comparison of the respective merits. Thirdly, a description of each technique in one volume may permit its adoption as an introductory manual for the administration, recording, and interpretation of many of these instruments. It is agreed by all that a text for the training of graduate students in this area is already sorely needed. Fourthly, the volume is being presented in order to facilitate much needed research on these tools, especially in the study of validity and reliability. The final purpose of the author is to stimulate the invention of new methods.

Chapter I contains definitions of projection, general characteristics of projective techniques, some theoretical foundations, and sources of basic theorems underlying these methods. This last section refers to the contributions of Freud, Gestalt theory, clinical psychiatry, cultural anthropology, and learning theory and considers how all of these influences converge in a modern dynamic theory of personality.

The rest of the book is divided into five parts, as follows:

Part I discusses word association and related techniques, in addition to incomplete sentences, the tautophone, and story telling and completion procedures.

Part II consists of the visual stimulus techniques such as Rorschach, cloud pictures, thematic apperception test and its modifications, picture-frustration study and the Szondi test.

Part III is a treatment of expressive movement and related techniques and concerns the work of Allport and Vernon, Werner Wolff, the analysis of handwriting, Mira myokinetic psychodiagnosis, visual motor tests, drawing and painting, finger painting, completing pictures, the mosaic test, and analysis of voice and speech.

Part IV deals with play, drama, and related techniques.

Part V is an overview and evaluation of projective techniques, including the criteria for the adequacy of a projective technique, criticisms of projective

techniques, and research needs for the future consideration of a more scientific approach to psychodynamic measurement.

Following each chapter the author presents a comprehensive and thorough bibliography. Several bibliographies contain references to the foreign literature as well as to the American journals. The most outstanding bibliography consists of 798 references and covers 50 pages (these are approximately all of the references to the Rorschach technique for the past 27 years since publication of *Psychodiagnostiks*). These features, as well as the author's lucid style of writing, recommend this volume as a text in clinical methodology for graduate students and as a reference book for psychologists and other clinicians concerned with human personality.

ARTHUR WEIDER.

University of Louisville School of Medicine.

WOLFF, WERNER. *Diagrams of the unconscious*. New York: Grune & Stratton, 1948. Pp. xiv+423. \$8.00.

After a brief preface the book is divided into two parts containing seven chapters each. Part I gives a concise, though mostly non-critical, survey of the literature and introduces Wolff's ideas on how handwriting might be made to succumb eventually to experimental procedures. He believes he has discovered certain consistent "themes of personality" of the gestalt entity type which display the properties of form patterning plus sequential continuity. These themes gain their stability and properties by being anchored in the individual's neuro-muscular-physiological structure. As such, they are expressed in any *movements* of that structure. Handwriting, a product of some of those movements, when correctly analyzed depicts the individualistic characteristics of certain of those themes. Therefore, Wolff attempts to demonstrate by reports of experiments and analyses of illustrated signatures how writers employ the sub-properties of height, length, position and shape in organizing the form pattern and sequential continuity of their writing samples. The personality themes expressed are *dynamic symmetry* or proportional relationships, *rhythm* or periodicity of these characteristic properties, and *configuration* or the organization of the interrelationships of these different properties.

In my opinion: (1) his theoretical points are sufficiently sound to warrant investigation, (2) his individual signature analyses are instructive, but (3) his claims for experimental verification *cannot* be confirmed on the basis of the material included or referred to in this section. Wolff creates the unfortunate impression that he believes if he writes a statement with words of sufficient authoritative "ring," tosses in a few figures that have the appearance of acceptable size, the reader will accept statistical significance and substantiation of claimed relationships (pp. 58, 308). He uses eight pages (46-53), for example, in reporting one of

his studies where, due to an inaccurate statistical procedure (failing to combine plus-minus scores), a non-existent significant trend was found; further, the problem the study was concerned with is only *slightly* related to the point under discussion. Incidentally, the graphs shown there are displaced.

Wolff refers his "lay reader" to Part II in preference to Part I, feeling it contains "the material demonstrating the principles that determine everybody's form and expression in his graphic" productions.

Here he all but ignores his efforts in Part I and falls back upon the classical approach to handwriting analysis. Throughout these seven chapters he implies that:

1. Most writing is composed of figures whose form gestalt is sufficiently similar to convention to be usually recognizable.
2. Written figures are characteristically changed by the writer from the printed norm but usually in non-distorting ways.
3. These changes are consistent for the individual.
4. The characteristic properties of these changes are accepted by more than one individual to be correlated with certain moods, degrees of intensity of response, attitudes, social roles and adjustment mechanisms.

Wolff uses many of his previously reported matching experiments plus other literature in substantiating these implications. The themes of personality are dropped in favor of a system of multi-traits. Chapter XIII tabularizes his various graphic elements, his grosser groupings of these, his list of personality traits with which he believes they are correlated, and a recording blank for systematizing and simplifying the analysis. The procedure is demonstrated on several famous signatures in the last chapter.

Here, too, Wolff desires to appear experimentally sound. However, (a) he subjectively selects his trait names and does not define them, (b) the great majority of these appear undefinable descriptively or operationally, (c) the experimental work reported lends little support to his claims of correlations between traits and graphic elements, (d) there are no tables of norms, and (e) his only stated caution is for the analyzer to be certain that an interpretation is done on the basis of the configurational organization of the graphic elements. How much better it would have been for Wolff to (1) frankly state that the work is still based on subjective clinical judgment, (2) that he still does not have experimental procedures which permit empirical evaluation of what he believes to be true from observations, and (3) that with these limitations he offers his findings to the *professional* reader, inviting assistance in solving the multitudinal problems involved. In that vein, this reviewer feels that Wolff will have more success by dropping his personality traits, returning to his concept of themes, and employing characteristic, describable response patterns with measurable properties.

This ought to be a worthwhile book. In many ways it is. Yet its

obvious faults pervade so much of its material and influence, to the stage of inaccuracy, so many of its conclusions that one is hesitant to endorse it in spite of them. With adequate research many of its suggestions should lead to conclusive results. However, since many acceptable projective techniques are available, demonstration of considerable usefulness for this method is mandatory before others would be justified in spending the requisite research time and money.

FRED Y. BILLINGSLEA.

Tulane University.

HARRIS, NOEL G. (Ed.) *Modern trends in psychological medicine*. New York: Hoeber, 1948. Pp. xiv+450. \$10.00.

The broad scope of investigation allowed by the title of these collected articles permits its authors to inquire into any crook and cranny of clinical psychology, psychiatry, child psychology, sociology, industrial psychology, mental hygiene and rehabilitation. The book may be said to have no beginning and no end; still, the editor in the introduction states that "an unorthodox layout may produce stimulus to thought not only in psychiatrists, but also in other members of our profession as well as in educationists and all those dealing with the character formation of children." Why a premeditated lack of organization or structure should be "a stimulus to thought" is not explained.

The range of this behavioral fantasia can best be understood by a list of the chapter titles in their order—The relation of psychological medicine to general medicine, the physiology of emotion, the importance of constitutional factors, the causative factors in mental disturbances, electrophysiology in psychiatry, diagnostic measures, marriage and the family, the future of child guidance, psychopathic personality, psychotherapy, recent technique of physical treatment and its results, further developments in abreaction, modern social and group therapy, principles of mental hygiene, character formation in relation to education, personnel selection, mental hygiene in industry, rehabilitation and the individual, psychological medicine and world affairs.

The first two chapters ("Psychosomatic medicine" and "The physiology of emotion") by N. G. Harris and S. Wright, respectively, present a brief but systematic survey of work that has been done in these fields, both here and in England, extending from approximately the early work of Cannon up to 1947. R. G. McInnes's chapter on "The causative factors in mental disturbances," consists of a well organized survey of the following factors—The genetic, constitutional, endocrine, psychological and social. F. Golla's article on electrophysiology in psychiatry and M. Jame's chapter on diagnostic measures have similar textbook value.

H. Stalker gives a sympathetic presentation of Henderson's theory concerning the nature of the psychopathic personality, but he concludes that the "key to the psychopathology of the psychopathic states is the

concept of immaturity." This emotional immaturity is not explainable in terms of environment. Stalker rather suggests that it is "an inherent defect, meaning that the capacity to mature is to a greater or lesser degree wanting." This seems to be a sensible variation of Prichard's notion that the psychopath is a "moral imbecile."

The chapters on psychotherapy, recent technique of physical treatment, modern social and group therapy, personnel selection and mental hygiene in industry are all competent surveys of these fields.

The most interesting chapter in the book is Jules Masserman's discussion of psychological medicine and world affairs. His entire argument is based on his biodynamic principles of behavior, which include—the principles of motivation, environment interpretation, substitutive adaptation and neurotogenic conflict. Concerning the principle of motivation, he says "The behavior of all organisms . . . is directed, in its deepest determinants towards the satisfaction of physiological needs." He leaves no room in his system for the existence of *acquired needs* which are not specifically physiological. But in 27 pages he attempts to explain national and internal problems and to suggest how most of them may be solved. His technique is unquestionably skillful. He refers to interesting analogies within the realm of his cat experiments in applying each of his four principles to the world at large. He, of course, protects himself each time, by reminding the reader that there is a difference between cats and men, but at the same time entices one to speculate concerning the importance of the existing similarities.

LIVINGSTON WELCH.

Hunter College.

BOOKS AND MATERIALS RECEIVED

- ANASTASI, ANNE, AND FOLEY, JOHN P., JR. *Differential psychology*. (Rev. Ed.) New York: Macmillan, 1949. Pp. ix+894. \$5.50.
- BRITT, STEUART HENDERSON. *Social psychology of modern life*. (Rev. Ed.) New York: Rinehart, 1949. Pp. xvi+703. \$4.50.
- BURROW, TRIGANT. *The neurosis of man*. New York: Harcourt, Brace, 1949. Pp. xxvi+428. \$7.50.
- CHAPANIS, A., GARNER, W. R., AND MORGAN, C. T. *Applied experimental psychology*. New York: John Wiley, 1949. Pp. xi+434. \$4.50.
- CRONBACH, LEE J. *Essentials of psychological testing*. New York: Harper, 1949. Pp. xiii+460. \$4.50.
- DASHIELL, J. F. *Fundamentals of general psychology*. (3rd Ed.) New York: Houghton Mifflin, 1949. Pp. x+690. \$4.00.
- FLESCH, RUDOLF. *The art of readable writing*. New York: Harper, 1949. Pp. xiv+237. \$3.00.
- FRANK, JEROME. *Courts on trial*. Princeton: Princeton Univ. Press, 1949. Pp. xii+441. \$5.00.
- GOODENOUGH, FLORENCE L. *Mental testing*. New York: Rinehart, 1949. Pp. xi+609. \$5.00.
- HILLPERN, ELSE P., HILLPERN, ED. P., AND SPAULDING, I. A. *Bristow Rogers: American Negro*. New York: Hermitage House, 1949. Pp. 200. \$3.00.
- HOPPOCK, ROBERT. *Group guidance*. New York: McGraw-Hill, 1949. Pp. 393. \$3.75.
- KARPMAN, BEN. *Objective psychotherapy*. Monogr. Suppl. No. 6, *J. clin. Psychol.*, July 1948. Pp. 154.
- LAWRENCE, MERLE. *Studies in human behavior*. Princeton: Princeton Univ. Press, 1949. Pp. x+184. \$3.50.
- LUNDHOLM, HELGE. *God's failure or man's folly?* Cambridge, Mass.: Sci-Art, 1949. Pp. 471. \$6.75.
- MEAD, MARGARET. *Male and female*. New York: William Morrow, 1949. Pp. 477. \$5.00.
- O'KELLY, LAWRENCE I. *Introduction to psychopathology*. New York: Prentice-Hall, 1949. Pp. xxi+733.
- PICHOT, PIERRE. *Les tests mentaux en psychiatrie*. Paris: Presses Universitaires de France, 1949. Pp. 238.
- POSTMAN, LEO, AND EGAN, JAMES P. *Experimental psychology*. New York: Harper, 1949. Pp. xiv+520. \$4.50.
- POTTER, MURIEL CATHERINE. *Reception of symbol orientation and early reading success*. Teach. Coll. Contr. Educ. No. 939. New York: Bureau of Publications, Teach. Coll., Columbia Univ., 1949. Pp. viii+69. \$2.10.

SALTER, ANDREW. *Conditioned reflex therapy*. New York: Creative Age Press, 1949. Pp. x+359. \$3.75.

SARGENT, S. S., AND SMITH, M. W. (ED.) *Culture and personality*. New York: Viking Fund, 1949. Pp. 219. \$1.50.

SHELDON, WILLIAM H. *Varieties of delinquent youth*. New York: Harper, 1949. Pp. xvii+888. \$8.00.

SIMPSON, GEORGE G. *The meaning of evolution*. New Haven: Yale Univ. Press, 1949. Pp. xv+348. \$3.75.

STONE, CALVIN P. *Case histories in abnormal psychology*. Stanford: Stanford Univ. Press, 1949. Pp. x+105. \$1.75.

THORNDIKE, E. L. *Selected writings from a connectionist's psychology*. New York: Appleton-Century-Crofts, 1949. Pp. vii+370. \$3.50.

VALENTINE, W. L., AND WICKENS, D. D. *Experimental foundations of general psychology*. (3rd Ed.) New York: Rinehart, 1949. Pp. xvii+472. \$3.00.

VAUGHAN, ELIZABETH HEAD. *Community under stress*. Princeton: Princeton Univ. Press, 1949. Pp. 160. \$2.50.

VERNON, P. E. *Personnel selection in the British Forces*. London: Little Paul's House, 1949. Pp. 324.

WALLIN, J. E. W. *Children with mental and physical handicaps*. New York: Prentice Hall, 1949. Pp. xxii+549. \$5.00.

WEITZMAN, ELLIS, AND MCNAMARA, W. J. *Constructive classroom examinations*. Chicago: Science Research Associates, 1949. Pp. xvi+153.

WITTKOWER, ERIC. *A psychiatrist looks at tuberculosis*. London: National Ass. for the Prevention of Tuberculosis, 1949. Pp. 151.

WOLLNER, MARY HAYDEN. *Children's voluntary reading*. Teach. Coll. Contr. Educ. No. 944. New York: Bureau of Publications, Teach. Coll., Columbia Univ., 1949. Pp. vii+117. \$2.35.

YOUNG, KIMBALL. *Sociology*. (2nd Ed.) New York: American Book Co., 1949. Pp. viii+638. \$5.00.

ZUCKER, CONRAD. *Psychologie des Aberglaubens*. New York: Grune & Stratton, 1949. Pp. 330. \$5.50.

A critical review of the research on elementary school curriculum organization 1890-1949. Urbana: Univ. of Illinois Bulletin, 1949. Pp. 29.

Educacion. No. 7-8. Lima, Peru: Universidad Nacional de San Marcos, 1948. Pp. 261.

Guidance worker's preparation. A directory of the guidance offerings of colleges and universities. Misc. 3333. Washington, D. C.: Federal Security Agency, Office of Education, 1949. Pp. 45.

ative

ality.

York:

Yale

ford:

ology.

tions

Pp.

eton:

ndon:

icaps.

room

-153.

ndon:

each.

each.

erican

York:

gani-

9.

e San

rings

ederal